26th International Congress of Mechanical Engineering
COBEM 2021
ABCM Associação Brasileira de Engenharia e Ciências Mecânicas
November 22-26, 2021, Virtual Congress, Brazil

COB-2021-1152
VERIFYING EXPLAINABILITY OF STEAM GENERATOR EFFICIENCY PREDICTION WITH SHAP VALUE INTERPRETATION

Lara Werncke Vieira
Augusto Delavald Marques
Jéssica Duarte
Rafael Petri Zanardo
Rodrigo Donni
Paulo Smith Schneider
Federal University of Rio Grande do Sul
lara.vieira@ufrgs.br
augusto.marques@ufrgs.br
jessica.jd.duarte@gmail.com
rafael.zanardo@hotmail.com
rodrigo.donni@gmail.com
pss@mecanica.ufrgs.br

Abstract. *Complex engineering systems, such as power plants, deliver their best performance when operating along a designed range of some priority parameters. Predicting a power plant efficiency requires models that are more flexible and thus more adaptable to the complex behavior of the real world, such as non-linear relationships and interactions between variables. Machine Learning (ML) is usually seen as an option for powerful algorithms with high accuracy but without intelligibility. The goal of this paper is to predict the steam generator efficiency of a coal-fired power plant and quantify the average contribution that each feature brings to the prediction made by the model. In this regard, Deep Learning (DL) techniques are applied to steam generation efficiency prediction, while Shapley Additive Explantions (SHAP) game theoretic approach was used to explain the outputs of the data-driven model. As a result, it quantifies the impact of each model input and analyze the model decisions, helping to build better models and guide the operators decisions. The impact of varying parameters that influence steam generator efficiency plays a vital role in the daily operational management of power systems without the need to choose between accuracy and explainability on the models.*

Keywords: *Artificial Neural Networks, Coal-fired power plant, Explainability, SHAP value interpretation.*

1. INTRODUCTION

The concept of explainable artificial intelligence (AI) has become more important with its gain of popularity. In many cases, AI is considered a "black box" due to its lack of explainability and transparency. In other words, we do not know exactly what is happening within the model. Researches had attempted to explain AI systems and various explainable techniques have been developed. The growing availability of big data has increased the benefits of using complex models, even so simple models may be preferred for easy interpretation. The trade-off between accuracy and interpretability reflect the ability to correctly interpret a prediction model's output (Kuzlu *et al.*, 2020; Lundberg and Lee, 2017b; Miller, 2019).

Interpretability is important for machine learning (ML) models because of three main points: trust building, human curiosity and learning, and detect biases. The machine learning model predictions is more reliable if the model can explain why it made the prediction while can facilitate learning and help it's users develop better understanding and intuition on the prediction problem. Such knowledge is an insight into how a model may be improved and point out the relevant parameters of the application field. The higher the interpretability of a machine learning model, the easier it is for someone to comprehend why a certain prediction was made by the model (Lundberg and Lee, 2017b; Molnar, 2021).

Interpretability can help in detecting biases at the time of model development. Interpretability can also help in identifying edge cases, where a model might fail. Determine the influence that the input parameters have on the output response in a system while analyzing the individual effects of each parameter and the interactions between them allows the development of a model that relates the response to the significative input parameters. This model can be used for improvements and support decision making (Kuzlu *et al.*, 2020; Molnar, 2021).

This paper applies explainable AI techniques to interpret the steam generator efficiency prediction based on machine learning method. The system in analysis is the steam generator of a coal-fired power plant. The algorithm was analyzed in terms of feature importance and parameter contributions. The results showed that efficiency can be improved by selecting high-importance features in the SHAP analysis, facilitating decision-making during plant operation.

2. SYSTEM DESCRIPTION

The system in analysis is the steam generator of the PECEM power plant. The PECEM is located near the ocean coast of the State of Ceará, Brazil, composed of three identical and independent power groups. Each group is designed to produce 360 MW out of Colombian coal with a lower heating value (LHV) about 25,750 kJ/kg, burned on a sub-critical steam generator. The furnace operates under balanced drought conditions; with natural circulation and steam reheat. A parallel back end splits flue gas flows through the primary superheater and the reheater exchangers (EDP, 2019; The Babcock and Wilcox Company, 2015). A schematic layout of the steam generator and its coupled coal mills is presented in Fig. 1.

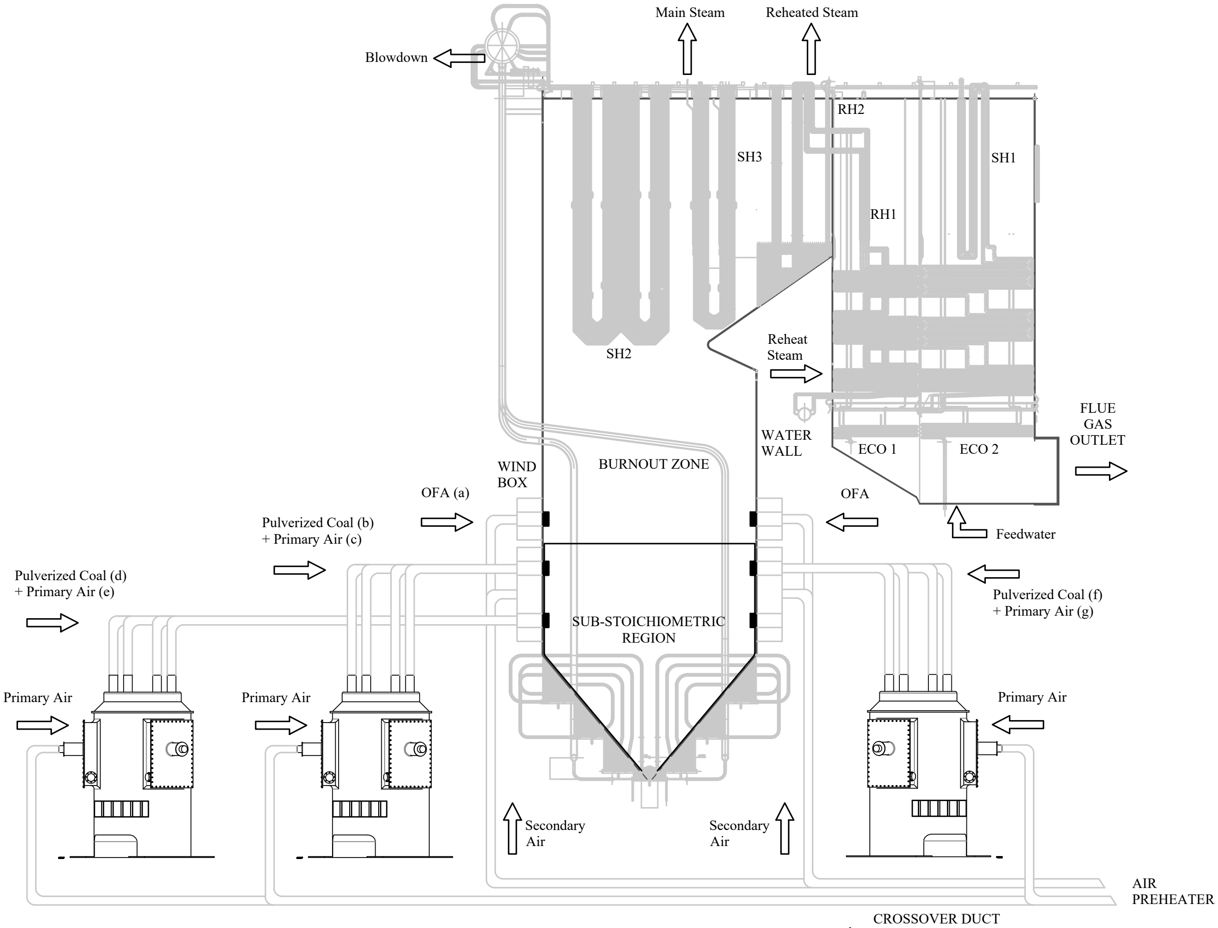


Figure 1: Steam generator schematic layout (UTE PECÉM, Brazil)

Preheated air stream coming from an external heat recovery device at approximately 300°C is split into two feeding paths, the primary and secondary air flows. Primary air is admitted in the mill to both perform coal drying and transport it to the steam generator burners. Each mill feeds a burner line of six pulverized coal combustors or burners, placed in independent wind boxes. The pulverized fuel and the primary air are introduced into the furnace via a combination of twenty four Low NOx Axial Swirl Burners (letters b to g in Figure 1) according to the load level, under sub-stoichiometric conditions. Combustion is completed on the furnace upper zone by twelve over fire air ports (OFAs, ports a in Figure 1).The feedwater arrives at 276°C and 168 bara, the output superheated steam at 538°C drives the vapour cycle The Babcock and Wilcox Company (2015).

3. METHODOLOGY

Figure 2 presents a flowchart of the proposed methodology.

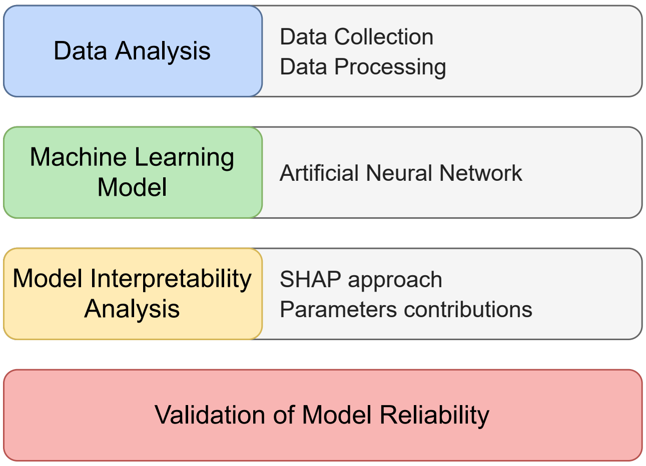


Figure 2: Steps of the proposed methodology

The data analysis procedure is dedicated to data collection and data processing. The data was acquired directly from the PECEM power plant Distributed Control System (DCS). Data processing identified the existence of special patterns, outliers, variation, and distribution. The input parameters were selected based on their controllability, which means they could be directly impacted by the actions of the unit control operator.

The selected ML model was Artificial Neural Networks (ANN) because of their ability to deal with complex problems. The metrics to evaluate the ANNs configuration performance were the Mean Absolute Error (MAE) and the Mean Squared Error (MSE), according to Equations (1) and (2).

$$MAE = \left(\frac{1}{n} \sum_{i=1}^{n} |X_{est} - X_{obs}| \right) \tag{1}$$

with X the steam generator efficiency for both measured (obs) and estimated (est) values and n the number of data points.

$$MAE = \left(\frac{1}{n} \sum_{i=1}^{n} (X_{est} - X_{obs})^2 \right) \tag{2}$$

MSE penalizes large errors values, while the MAE evaluates the absolute error, without differentiating individual error weights.

ANN gathers information from the environment through data. The processed dataset was divided into training and test group to evaluate the best set configuration. All set configurations were used to build different ANNs, evaluated by their MAE and MSE in order to identify the best set configuration. The hyperparameters (number of hidden layers, number of hidden neurons per each hidden layer, and activation functions) were defined through an iterative approach starting with the simplest ANN with a single hidden layer. ANN Multi-Layer Perceptron (MLP) type was chosen to predict the steam generator efficiency. The list of variables considered in the model is presented in Table 1.

Table 1: Summary of model variables

Variable	Unit	Description
P1	kg/s	Primary Air Flow
P2	C	Pulverized coal outlet temperature
P3	rpm	Speed of the dynamic classifier
P4	kg/s	Secondary Air Flow
P5	%	Excess O_2
P6	mbar	Secondary air pressure
P7	mbar	Primary air pressure
S1	%	Steam generator efficiency

The model interpretability analysis was made to evaluate the contributions of the variables through the SHAP approach. SHAP (SHapley Additive exPlanation) is a game theoretic approach to explain the output of any machine learning model. It connects optimal credit allocation with local explanations using the classic Shapley values from game theory and their related extensions. The Shapley value is the average marginal contribution of a feature value across all possible coalitions (Lundberg and Lee, 2017b; Lundberg, 2018).

Equation (3) computes the Shapley value ϕ_i for a given feature i (Lundberg and Lee, 2017a).

$$\phi_i = \sum_{S \subseteq F/\{i\}} \frac{|S|!(|F|-|S|-1)!}{F!}[f_{S\cup\{i\}}(x_{S\cup\{i\}}) - f_S(x_S)] \tag{3}$$

where F is the set of all features and S is a subset of F that in each sum contains all features but i.

Among SHAP contributions, it includes new methods for computing feature importance values with improved computational performance and better consistency with human intuition. The goal is to apply an unified framework for interpreting predictions and validate the model reliability.

4. RESULTS

The power plant Distributed Control System (DCS) continuously acquired the half-hour mean values of the parameters data during operation from August 2018 to October 2019. The only parameter in this group that is not directly measured on site is the steam generator efficiency that is calculated through other secondary parameter measures available.

The dataset was analyzed and processed to remove gross errors and outliers such as negative and null observations. The data was also filtered by electric power generation to reflect the 340 to 365 MW range. The 3048 samples was randomized and split into 70% training and 30% for testing and validation. The ANNs were developed using Python programming language through Keras programming interface (Chollet and al, 2015) on top of the Tensorflow machine learning library (Abadi *et al.*, 2015).

The activation functions included ReLU (Rectified Linear Unit) and Tanh (hyperbolic tangent). The investigation process aimed at the simplest ANN capable of representing the system behavior. The chosen topology and results for the ANN are presented in Figure 3 and in Table 2.

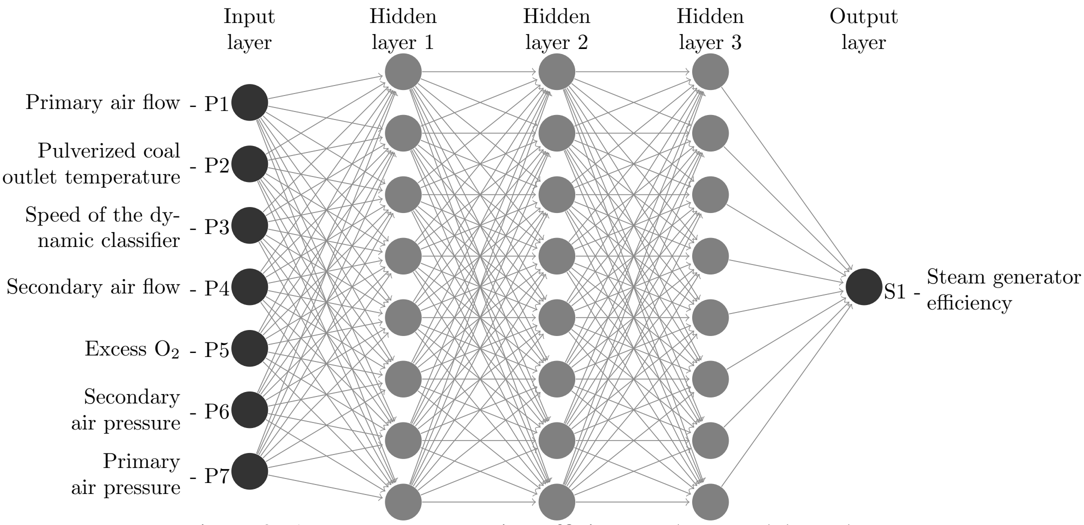


Figure 3: ANN steam generation efficiency subset model topology.

Table 2: Chosen ANN - Backpropagation learning algorithm and Multi-Layer Perceptron network type for 200 epochs with a batch size of 256.

Input layer	Hidden neurons	Hidden layers	Activation function	Output layer	MAE training	MSE training	MAE testing	MSE testing
7	8 - 8 - 8	3	tanh-tanh-relu	1	0.0103	0.0002	0.0106	0.0002

The inputs and output parameters, hidden layer and the number of neurons are presented in Figure 3. The ANN topology selection was evaluated by trial and error by changing the number of neurons and the consequent MSE and MAE values for each individual ANN. The output layer with the steam generator efficiency in this architecture holds one only neuron.

The SHAP approach was applied to evaluate the feature importance and the respective contribution or impact of each of the 7 variables on the model output. The graphical result is presented in Figure 4.

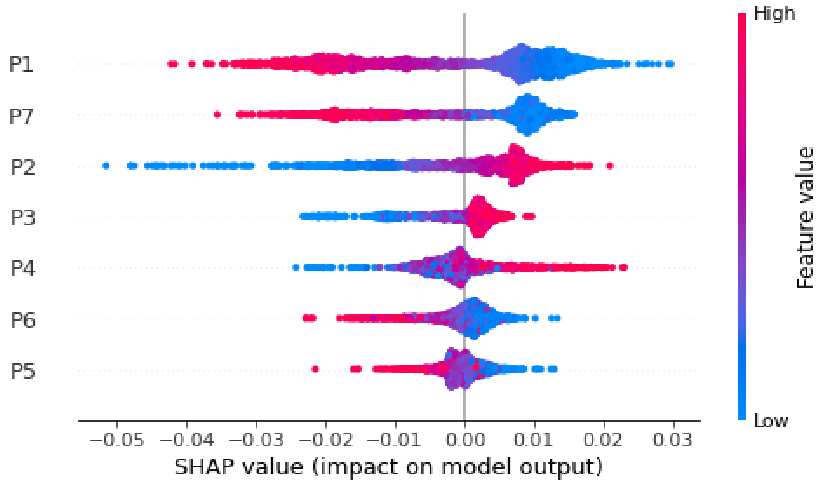
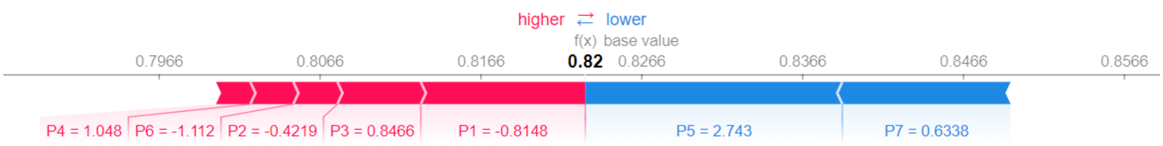


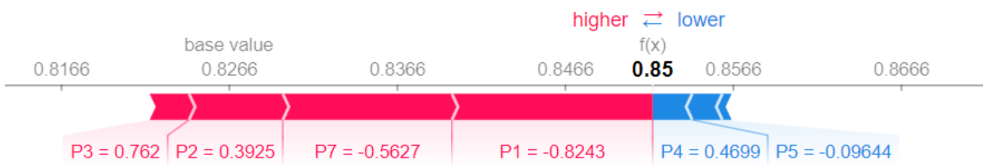
Figure 4: SHAP values with impact on model

The higher the value of the feature, the more reddish its color is on the graphic. In the features with a clear impact on the developed ANN model, the red dots and the blue dots are clearly splitted on the resulting SHAP values axis (horizontal axis), indicating whether the output varies according to the input variables. The greater the distribution of points along the SHAP values axis, the greater the impact of the feature on the output. If the colors are mixed in the centerline, the impact of the feature is not that representative.

In the case of primary air flow (P1), higher feature values indicate negative SHAP values. High values of features with this behaviour are actually helping raise the chance of a lower steam generator efficiency. The impact of each feature in the output result for two arbitrary conditions is explored in Figures 5a and 5b.



(a) Prediction for a condition with resulting steam generator efficiency of 82%



(b) Prediction for a condition with resulting steam generator efficiency of 85%

Figure 5: SHAP explanation for the steam generator efficiency output for two different conditions

.

Each plot displays the base probability value 82,66%, over the training dataset. The impact of each feature is represented in blue or red color. Blue color indicates that the feature caused a decrease in the output value, and the red color indicates that the feature value caused an increase in the output value. So, in the plots presented in Figure 5a and 5b, the "red" features are actually helping raise the chance of increasing steam generator efficiency, while the negative features are lowering the chance. The results for these conditions are in accordance with the SHAP value with impact on model output, regarding Figure 4. In both cases, the primary air flow (P1) is the most important variable regarding impact on the output.

5. CONCLUSION

This paper presented the application of ANNs models to predict the steam generator efficiency of a coal-fired power plant. The model was based on acquired data from the PECEM power plant, Brazil. The identification of feature importance its essential to validation of model reliability while in real life applications the operators of the power plant can concentrate their efforts to control the really impacting variables.

The best ANN architecture for seven-input combination was made of three hidden layers with eight neurons in the hidden layers and one neuron in the last layer corresponding to the model output. The training dateset presented MAE and MSE of 0.0103 and 0.0002 and the testing dataset 0.0106 and 0.0002, respectively. The application of SHAP allowed to identify the impact of each input feature on model output and respective importance on prediction. The most important feature was the primary air flow (P1), followed by primary air pressure (P7), pulverized coal outlet temperature (P2), speed of the dynamic classifier (P3), secondary air flow (P4), secondary air pressure (P6) and excess O_2 (P5). The impact of varying parameters that influence steam generator efficiency plays a vital role in the daily operational management of power systems without the need to choose between accuracy and explainability on the models.

6. ACKNOWLEDGEMENTS

Authors acknowledge Energy of Portugal EDP for the financial and technical support to this project; L.W.Vieira acknowledges the financial support from CAPES 88882.346360/2014-01 for her Ph.D. grant; A.D.Marques acknowledges the financial support from CNPq 132422/2020-4 for his MSc grant; J.Duarte acknowledges the financial support from CNPq 154147/2020-6 for her undergraduate scholarship; P.S. Schneider acknowledges CNPq for his research grant (PQ 305357/2013-1).

7. REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viegas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y. and Zheng, X., 2015. "TensorFlow: Large-scale machine learning on heterogeneous systems". URL `https://www.tensorflow.org/`.

Chollet, F. and al, E., 2015. "Keras". URL `https://keras.io/`.

EDP, 2019. "UTE PECÉM".

Kuzlu, M., Cali, U., Sharma, V. and Güler, Ö., 2020. "Gaining insight into solar photovoltaic power generation forecasting utilizing explainable artificial intelligence tools". *IEEE Access*, Vol. 8, pp. 187814–187823. ISSN 21693536. doi: 10.1109/ACCESS.2020.3031477.

Lundberg, S., 2018. "SHAP documentation". URL `https://shap.readthedocs.io/en/latest/`.

Lundberg, S. and Lee, S.I., 2017a. "A unified approach to interpreting model predictions". *arXiv preprint arXiv:1705.07874*.

Lundberg, S.M. and Lee, S.I., 2017b. "A unified approach to interpreting model predictions". *Advances in Neural Information Processing Systems*, Vol. 2017-Decem, No. Section 2, pp. 4766–4775. ISSN 10495258.

Miller, T., 2019. "Explanation in artificial intelligence: Insights from the social sciences". *Artificial Intelligence*, Vol. 267, pp. 1–38. ISSN 00043702. doi:10.1016/j.artint.2018.07.007. URL `https://doi.org/10.1016/j.artint.2018.07.007`.

Molnar, C., 2021. "Interpretable Machine Learning". URL `https://christophm.github.io/interpretable-ml-book/`.

The Babcock and Wilcox Company, 2015. *Steam: its generation and use*. The Babcock and Wilcox Company, 42nd edition. ISBN 978-0-9634570-2-8.

8. RESPONSIBILITY NOTICE

The authors are solely responsible for the printed material included in this paper.