ANALYSIS OF INFLUENCE OF CHEMICAL COMPONENTS IN THE MECHANICAL PROPERTIES IN THE STEEL DATABASES USING MACHINE LEARNING

Conference Paper · January 2022

CITATIONS
0

READS
28

5 authors, including:

Ana Costa
University of Porto
10 PUBLICATIONS **16** CITATIONS
SEE PROFILE

Joao Manuel R. S. Tavares
University of Porto
1,035 PUBLICATIONS **12,532** CITATIONS
SEE PROFILE

J. M. A. Cesar de Sa
University of Porto
139 PUBLICATIONS **2,248** CITATIONS
SEE PROFILE

Congreso de Métodos Numéricos en Ingeniería
12-14 septiembre 2022, Las Palmas de Gran Canaria, España

ANALYSIS OF INFLUENCE OF CHEMICAL COMPONENTS IN THE MECHANICAL PROPERTIES IN THE STEEL DATABASES USING MACHINE LEARNING

Ana P. O. Costa[1], Mariana R. R. Seabra[1,2], João Manuel R. S. Tavares[1,2], José M. A. César de Sá[1,2] and Abel D. Santos[1,2]

[1] Departamento de Engenharia Mecânica, Faculdade de Engenharia, Universidade do Porto
s/n, R. Dr. Roberto Frias, 4200-465 Porto, Portugal
E-mail: anapauladeoliveirac@gmail.com page: http://www.fe.up.pt

[2] Instituto de Ciência e Inovação em Engenharia Mecânica e Engenharia Industrial (INEGI)
Campus da FEUP, R. Dr. Roberto Frias 400, 4200-465 Porto
web: http://www.inegi.pt

Keywords: Feature Selection, Feature importance, Data-Driven, Material Design

Abstract. *Feature selection is an important step to reduce the complexity of a dataset and decrease the required computational time to classify a sample. There are plenty of methods used in Machine Learning (ML) serially falling into the categories of supervised and unsupervised. Some of the commonly known unsupervised methods are K-mean, the simplest one, and the random forest, more complex, and both are capable to find a pattern within a dataset allowing the clustering of similar data. For material design using artificial intelligence, it is crucial to have data selection. In steel databases, there are many features that may be relevant in the process of designing a new alloy, such as chemical composition, temper, cooling curves, or microstructures. However, when accounting for all these features, the complexity of the solution and computational time may increase. Hence, in this work, as a first step to design a new alloy using ML, the influence of chemical components are analyzed, in order to define which are more significant to determine the mechanical properties, and reduce the number of used features as input in a neural network, for example. Thus, Machine Learning techniques are explored in this context.*

Acknowledgements Authors gratefully acknowledge the funding of Project NORTE-01-0145-FEDER-032419, co-financed by "Programa Operacional Regional do Norte" (NORTE2020), through "Fundo Europeu de Desenvolvimento Regional" (FEDER) and by "Fundação para a Ciência e Tecnologia" through its component of the state budget. Also, the first author greatly acknowledges the funding of the doctoral grant 2021.05067.BD by "Fundação para a Ciência e Tecnologia" (FCT), in Portugal, through the national budget and Community budget from European Social Fund (ESF).

1 INTRODUCTION

For long decades, the design of new alloys has faced challenges and costly try and error methods. These methods require a full characterization of the physical and mechanical characteristics of the new material [1] [2]. As a consequence, there is a great amount of data available in the field of alloy metallurgy. Steel industry is one of the oldest and the biggest in the field of metallurgy, having produced a wide range of alloys, with good parameters for strength, corrosion, and many others. However, even these days, industries are still looking for designing new alloys to attend to a variety of demands [3] [4] [5] [6]. There are many elements that can be combined with different weights to produce new alloys. Nonetheless, in the try and error method, it is not affordable in cost and time to combine and produce all these possibilities and test which one is better.

The development of computational tools brings the possibility of solving this issue through numerical methods such as Machine Learning (ML). ML has been used to solve complex problems, requiring a database with previous information. The main task of the ML is trying to extract patterns in input data [7] [8]. There are many algorithms in ML, which can be divided into regression, and classification algorithms. To use these methods, it is important to define an input and expected output. In some cases, the relevant dataset is huge and complex, making the process of choosing the inputs difficult and compromising the expected results. Additionally, the amount of data may increase computational cost and time. All these issues may be solved resorting to Feature selection.

Feature selection, also known as variable selection, are a group of ML algorithms used to understand the dataset and to reduce the number of input variables when developing a predictive model [9]. It may lead to a reduction of computational cost and to an improvement of the performance of the model. Moreover, Feature importance is a branch of feature selection, which refers to techniques to score the features input based on its importance to determine predicting values.

In this article, methods to measure the feature input importance are proposed for a Regression Random Forest (RF), and Fully Connected Neural Network (FCNN) algorithms thought the shapely addictive explanation (SHAP). Outcomes gotten from both algorithms were compared with each other. Finally, these results are used to reduce the input features, and their impact is evaluated.

2 METHODS

2.1 Random Forest

Random Forest derived from another popular ML method known as Decision Tree; as the name implies, this algorithm is made up of a decision tree, and where a target variable is predicted by learning simple decision rules inferred from data features. This algorithm is used in classifier and regression problems [10]. As in the Decision tree, the Random

Forest is also composed of tree of decisions, but, in this case, is not just one tree of decision, but many of them, and the algorithm test all those trees, and uses averaging to improve the predictive accuracy, control over-fitting, and have a final decision. To achieve the results, data are bagged in small groups and tested by the number of decision trees chosen, Figure 1 illustrates schematically the operation of this algorithm.

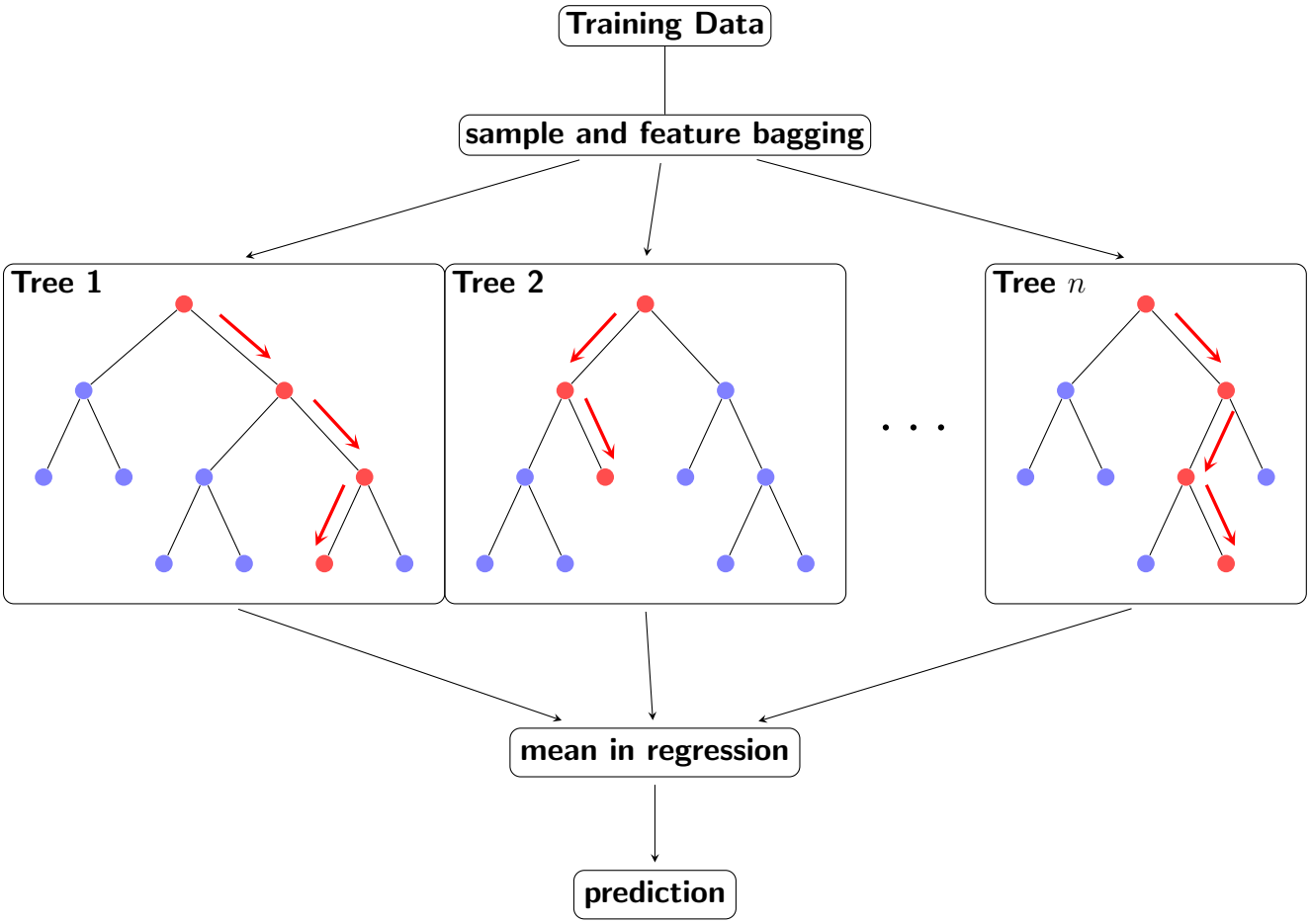

Figure 1. Scheme of a Random Forest classifier.

In mathematical terms, the Random Forest for regression is formed by growing trees depending on a random vector Θ such that the tree predictor $h(x,\Theta)$ takes on numerical values [11].

A random vector is created by a number of bagging randomly chosen instances in the training set. A random split selection is made up of a set of independent random integers ranging from 1 (one) to K. Its dimensions and character are determined by how it is used in tree construction.

Assuming that Y is the numerical expected outcome, and X is the input feature, for each numerical predictor $h(x, \Theta)$, the mean-squared generalization error is:

$$E_{\mathbf{X},Y}(Y - h(\mathbf{X}))^2 \tag{1}$$

where h(x,Θ_k) is a random forest predictor that is created by averaging across k trees. As the number of trees in the forest grows to infinity, almost certainly one has:

$$E_{\mathbf{X},Y}(Y - av_k h(\mathbf{X}, \Theta_k))^2 \rightarrow E_{\mathbf{X},Y}(Y - E_{\Theta} h(\mathbf{X}, \Theta_k))^2 \tag{2}$$

Also, it has to be assumed that for all Θ, $EY = E_X, Y(Y - h(X, \Theta)$. Then, one has:

$$E_{\Theta} E_{\mathbf{X},Y}(Y - av_k h(\mathbf{X}, \Theta_k))^2(forest) \leq \overline{\rho} E_{\Theta} E_{\mathbf{X},Y}(Y - av_k h(\mathbf{X}, \Theta_k))^2(tree) \tag{3}$$

where $\overline{\rho}$ is the weighted correlation between the residuals Y-h($\mathbf{X}$,Θ') where Θ, Θ' are independent.

RF is a popular method because of the possibility of interpreting the results, in spite of their non-linearity. Feature importance in RF is provided, after training, by computing the mean and the standard deviation of accumulation of the impurity, that is, the mean and the standard deviation of the measure of the chosen optimal condition. Impurity decreases within each tree [12]; thus, the higher the increment in purity, the higher will be the importance of the feature. This is done for each tree, then, it is averaged among all the trees and, finally, normalized to 1 (one). So, the sum of the importance scores calculated by a Random Forest is 1 (one).

2.2 Neural Network

Neural networks have become very popular in machine learning models. Usually, it is a supervised learning method that uses a set of classified training data to determine patterns relating to the inputs and outputs. It is compared with the human brain because it is constituted of layers of neurons or nodes that are capable to find relations between features in the data set, as displayed in Figure 2.

The disadvantage of this method is that it can not be easily interpreted, as it is difficult to trace how the neurons establish these relations; consequently, it is hard to evaluate the data set in order to reduce it, and/or improve the outcomes.

In the context of this work, the fully connected neural network is used. Hence, an overview of this method is presented in the following.

The FCNN has all neurons and layers connected, and this relation can be expressed by:

$$\mathsf{x}_i^k = \phi\left(\sum_j \mathsf{x}_j^{k-1} \mathsf{A}_{ij}^{k-1}\right) \tag{4}$$

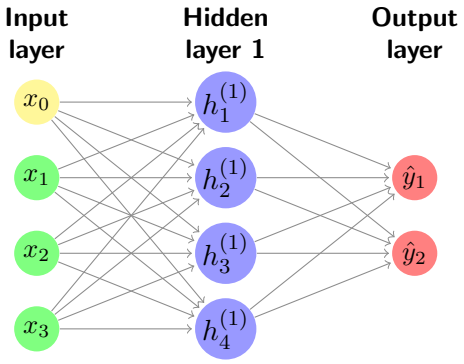

Figure 2. Scheme of a fully connected neural network.

where the x^k and x^{k-1} are vectors representing nodal values and bias of neurons in layers k and $k-1$, respectively, A_{ij}^{n-1} is the weight matrix, and $\phi(\mathsf{x})$ is the activation function. The activation function determines whether or not a neuron is activated, adding non-linearity to the NN. The Activation Function's main function is to convert the node's summed weighted input into an output value that may be passed to the next hidden layer or used as output [13]. There are many types of activation functions, and not a singular way to define the better to choice. After careful analysis of activation functions and their advantages, Relu was chosen due to its performance [14].
The Relu is defined as:

$$\phi(\mathsf{x}) = \begin{cases} \mathsf{x}\,, \ \mathsf{x} > 0 \\ 0\,, \ \mathsf{x} \leq 0 \end{cases} \tag{5}$$

The parameters that must be fitted during training are kept in the weight matrices $\mathsf{A}ij$, which, in this work, are fitted in a Gradient Descent method using the backpropagation algorithm [15]. In short, when using backpropagation, the NN's output is compared to the goal output, yielding an error function, $\mathsf{E}(\mathsf{w})$, which is dependent on the network's weights and biases. Then, using the following equation, weights and biases are updated:

$$\mathsf{w}^{t+1} = \mathsf{w}^t - \alpha \nabla_{\mathsf{w}} \mathsf{E}(\mathsf{w}^t) \tag{6}$$

where w^t and w^{t+1} are the vectors containing all the weights and biases of the NN, at time instants t and $t+1$, respectively, $\nabla_{\mathsf{w}}\mathsf{E}(\mathsf{w})$ is the gradient of the error function in respect to the weights and biases vector, and α is the learning rate, which, in this work was set to 0.03. Subsequently, a Loss function is used to evaluate the error. Smooth L1 Function was chosen in this study because, when compared to alternatives, it has a lower sensitivity to outliers and capacity to prevent explosive gradients. It is defined as:

$$\begin{cases} \frac{0.5(\mathsf{y}_i - \mathsf{y}_p)^2}{\delta} & if \ \ |\mathsf{y}_i - \mathsf{y}_p| < \delta \\ |\mathsf{y}_i - \mathsf{y}_p| - 0.5\delta \ , & otherwise \end{cases} \tag{7}$$

where y_i and y_p are the vectors containing the actual output ant the predicted output, and δ is a hyperparameter, which was set equal to the learning rate, $\delta = \alpha = 0.03$.

2.3 Shapley Additive Explanations

Shapley Additive Explanations (SHAP) is a unified way to explain the outcome of any deep learning model linking game theory with local explanations, bringing together many methodologies, such as a model-agnostic approximation approach (Kernel SHAP), and approximation methods (Gradient SHAP, Deep SHAP) [16]. It can be seen as an additive feature attribution method. In this type of methods, an effect is attributed to each feature, and the final explanation model is obtained by summing the effects of all feature attributions which approximate the output.

SHAP joins a local linear explanation model (Linear LIME), an additive feature attribution approach that employs a linear explanation model to locally approximate [17] that Shapley values that are the only viable solution to the objective function that LIME minimizes, and they meet the three conditions listed bellow, namely local accuracy, missingness, and consistency. Kernel SHAP is an additive feature attribution approach that also satisfies these three conditions. As a result of this union, Kernel SHAP improves the sample efficiency of SHAP value estimations without taking the model type into consideration.

Property 1 Local accuracy:
When approximating the original model f for a specific input x, local accuracy requires the explanation model g to at least match the output of f for the simplified input $x^{'}$:

$$f(x) = g(x^{'}) = \phi_0 + \sum_{i=1}^{M} \phi_i x_i^{'} \tag{8}$$

The explanation model $g(x^{'})$ matches the original model $f(x)$ when $x = h_x(x^{'})$, where $\phi_0 = f(h_x(0))$ represents the model output with all simplified inputs toggled off.
Property 2 Missingness:
A Feature missing in the original input is attributed no importance:

$$x^{'} = 0 \rightarrow \phi_i = 0 \tag{9}$$

Missingness constrains features where $x_i^{'} = 0$ to have no attributed impact.
Property 3 Consistency:
Consistency means that, even when updating a model to make a feature have a larger influence on it, the attribution ascribed to that feature will never diminish. Let $f_x(z^{'}) = f(h_x(z^{'}))$ and $z_{ii}^{'}$ denote setting $z_i^{'} = 0$. For any two models f and $f^{'}$, if

$$f_x^{'}(z^{'}) - f_x^{'}(z_{ii}^{'}) \geq f_x(z^{'}) - f_x(z_{ii}^{'}) \tag{10}$$

for all input $z^{'} \in \{0,1\}^M$, then $\phi_i(f^{'}, x) \geq \phi_i(f, x)$

Properties 1, 2, and 3 are classical Shapley values

SHAP values can be calculated, as unified measures of feature importance, by defining $f_x(s) = f(h_x(z^{'})) = E[f(x)|x_s]$ where S is the set of non-zero indexes in $z^{'}$ and $E[f(x)|x_s]$ is the expected value of the function conditioned on a subset S of input features. SHAP values combine these conditional expectations with game theory and with classical Shapley values to attribute ϕ_i values to each feature. Only one possible explanation of the model that follows Equation 11 and according [16], satisfies the three properties is as follows:

$$\phi_i = \sum_{S \subset N_{ii}} \frac{[S]!(M - [S] - 1)!}{M!}[f_x(S \cup i) - f_x(S))] \tag{11}$$

where N is the set of all input features.

2.4 Metrics

Metrics may be used to assess the performance of RF and NN classifiers. Three measures were used in this study, namely, Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and R-squared Score (R2 Score).

RMSE is given as:

$$\mathsf{RMSE} = \sqrt{\frac{\sum(\mathsf{y}_i - \mathsf{y}_p)^2}{n}} \tag{12}$$

where y_i and y_p are vectors representing the actual and projected outputs, respectively, and n is the number of samples. The root mean square error (RMSE) is a measure of how concentrated the data is around the best-fit line. MAE, on the other hand, is the average deviation of all samples:

$$\mathsf{MAE} = \frac{|\mathsf{y}_i - \mathsf{y}_p|}{n} \tag{13}$$

The R2 Score measures the percentage of correct predictions given by the model:

$$\mathsf{R2\ Score} = 1 - \frac{\sum(\mathsf{y}_i - \mathsf{y}_p)^2}{\sum(\mathsf{y}_i - \bar{\mathsf{y}}_i)^2} \tag{14}$$

where $\bar{\mathsf{y}}_\mathsf{i}$ is the mean of all real values. These three metrics are utilized to evaluate the RF and NN for an regression problem, and identify the best model, that can be used for feature prediction.

3 DATASET

The choice of the dataset may be determinant in successful model development for an ML algorithm. In this work, a database containing chemical composition, hardness, yield strength, and tensile strength for 70 different steel alloys was built. Due to the short size of the dataset, over-sampling was used in the training set, resulting in a total of 200 data points. The Data was extracted from *steel-grades.com* [18]. From the total number of data points, 80% were used for training. Subsequently, the remaining 20% of the data was shown to the test set NN and inferred values were compared with the actual values.

4 RESULTS AND DISCUSSION

It is already known that chemical composition plays an important role to determine the mechanical properties [19] [20] [21] [22] [23]; however, it is unknown which are the elements that are more important to determine those properties in machine learning models. A reduction the number of input the computational time and, additionally, it may improve the performance of the model.

Hence, in this work, two models were developed: Random Forest and FCNN, to analyze two related mechanical properties: Yield Strength(YS) and Tensile Strength(TS).

Random Forest

The first model developed, trained and tested is the Regression Random Forest for the YS. Fifty estimators are employed and K parameter related to random slip selection is set to 5. Figures 3 a, and b, present the model inference to the actual and predicted values, and the scatter plot, in a test set calculated for Yield Strength. This model was developed with the aim to verify which chemical components play a major role in the ML models.

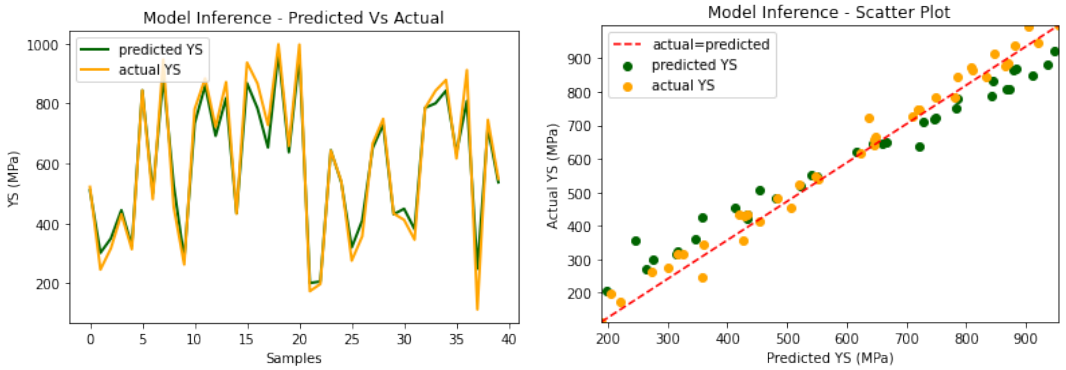

Figure 3. Yield Strength of different steel grades calculated by Random forest and respective tabled values.

The features that play a major role in the model RF-YS are exposed in Figure 4. In all models used in this work, hardness was included in the feature input together with chemical composition, as it lead to a better prediction [14]. In fact, adding the hardness as a input compensates the missing information in terms of the alloy fabrication, cooling curves, etc., [24] [25]. Moreover, hardness test are much simpler than most of experimental tests required to determine other mechanical properties.

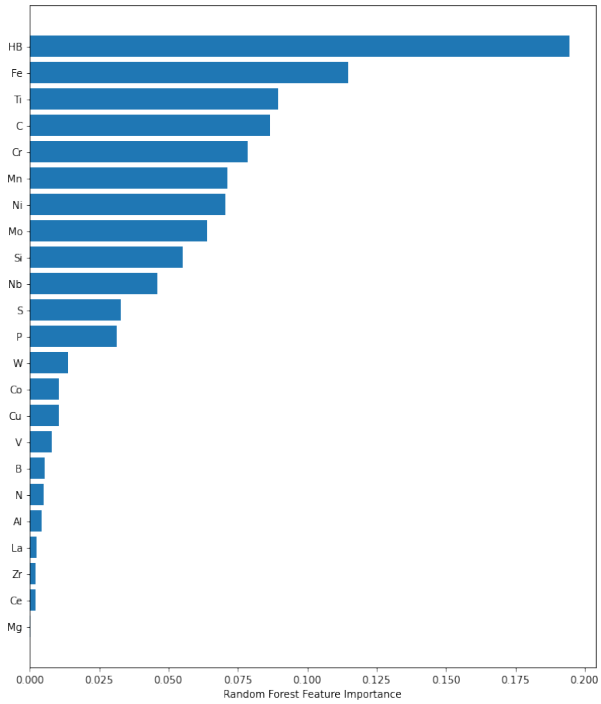

Figure 4. Random Forest Feature importance calculated for Yield Strength.

As a results, the three elements that have more impact in the case of YS besides hardness are: Fe, Ti, and C. Iron (Fe) is the most abundance element in steel alloys, which together with carbon that forms steel. Carbon can be presented in higher and low percentages, depending on desire properties; usually, the C adds hardness and strength through heat treatment. Titanium (Ti) is the element most usually added to the steel because it boosts strength and corrosion resistance of steel alloys. Ti also adds lightness, which, in general, is a desirable property, as it has a density less than half of density of steel [21] [22] [23].

The next model developed is used to determine the tensile strength through the RF algorithm. Figures 5a and b, present the model inference to the actual, and predicted values, and the scatter plot, in a test set. Table 1 presents values of MAE, MSE, and R2 found for this model.

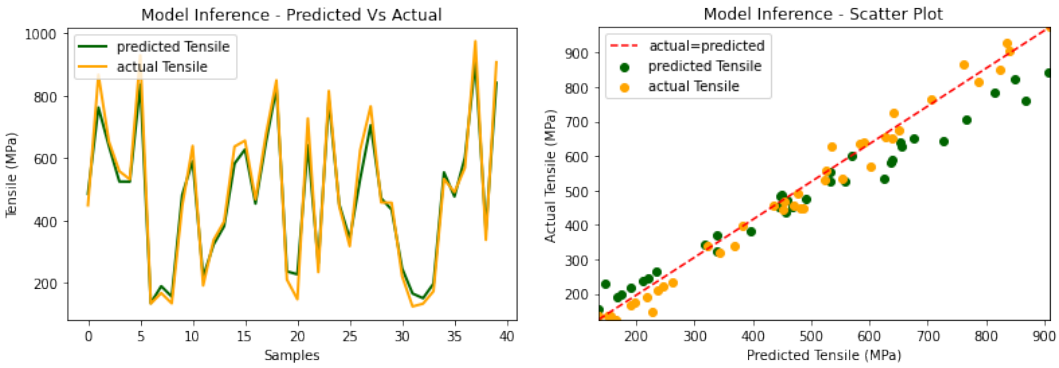

Figure 5. Tensile Strength of different steel grades calculated by Random forest and respective tabled values.

The features that play a major role in the model RF-TS are exposed in Figure 6.

The three elements that have more impact in the case of TS besides hardness, are the Fe, P, Cr, and, in this case, Fe plays a major role than hardness. P and Cr are usually added elements to increase the properties of Steel alloys; however, they are not essential as Fe, and C. Phosphorus increases strength and hardness at the expense of ductility and toughness, especially in quenched and tempered higher carbon steels. As a result, its concentration in most steel is capped at 0.05%. When used as an alloy in steel, phosphorus prevents light-gauge sheets from sticking. It enhances machinable in free-cutting steels, strengthens low carbon steel to some extent, and boosts corrosion resistance. When it comes to welding, phosphorus levels above 0.04% render the weld brittle and more likely to split. The molten weld metal's surface tension is reduced, making it difficult to manage. Chromium (Cr) It significantly improves the hardenability of steel and the corrosion resistance of alloys in oxidizing conditions. Its presence in some steels has the potential to induce excessive hardness and cracking in and around welds [21] [22] [23].

	MAE (MPa)	MSE (MPa)	R2
RF-YS	30.72	42.13	0.97
RF-TS	35.72	43.97	0.97

Table 1. Metrics found as to Random Forest.

Thought the metrics were possible to check that the RF presented a good performance, for both YS and TS.

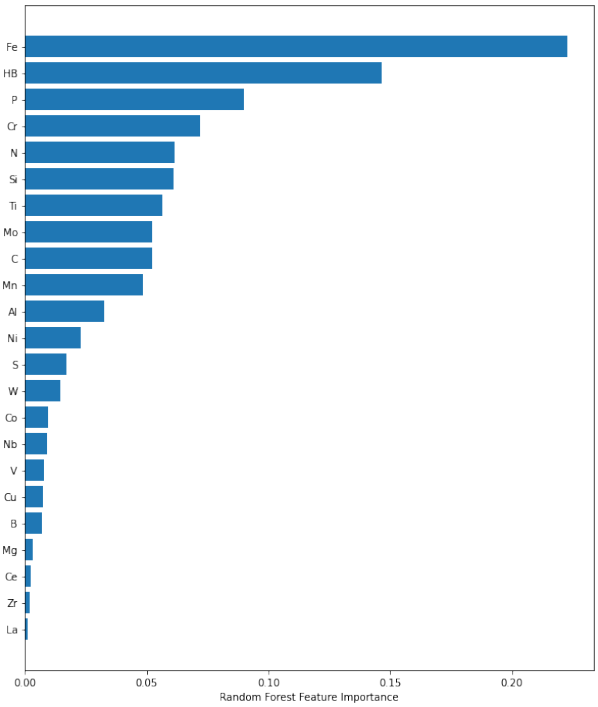

Figure 6. Random Forest Feature importance calculated for Tensile Strength.

FCNN

The second model developed, trained and tested is the FCNN for YS. Figures 7a and b, present training set and loss interference, and Figure 8a and b, present the model inference to the actual, and predicted values, and the scatter plot, in a test set calculated for Yield Strength. The results from the metrics and loss used to evaluate the model are presented in Table 4.

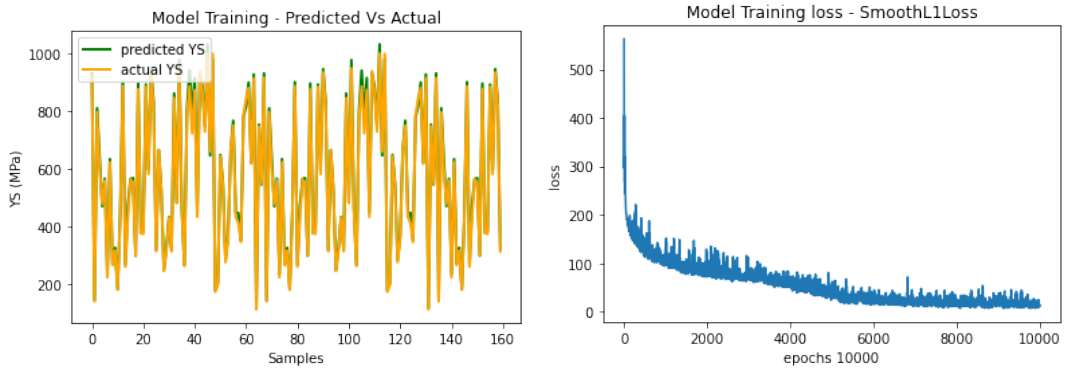

Figure 7. Training and Loss when calculating Yield Strength of different steel grades using Neural Network.

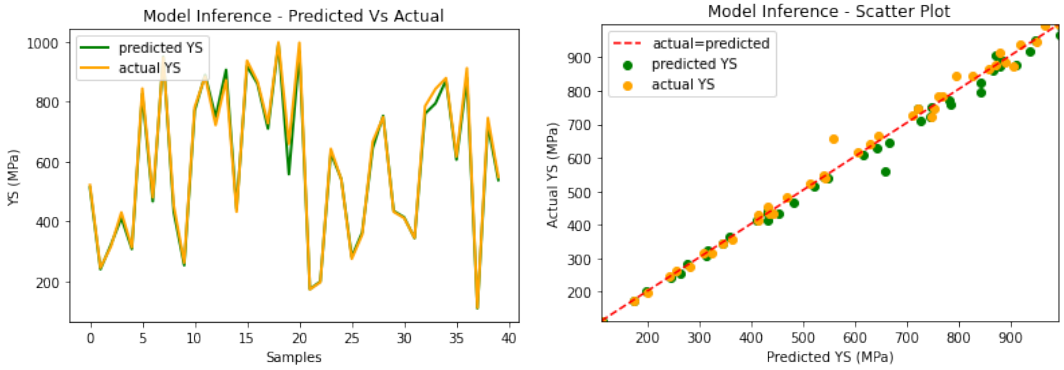

Figure 8. Test and Inference when calculating Yield Strength of different steel grades using Neural Network.

To analyze the feature importance, in this case, the SHAP function was used, and results are presented in Figure 9.

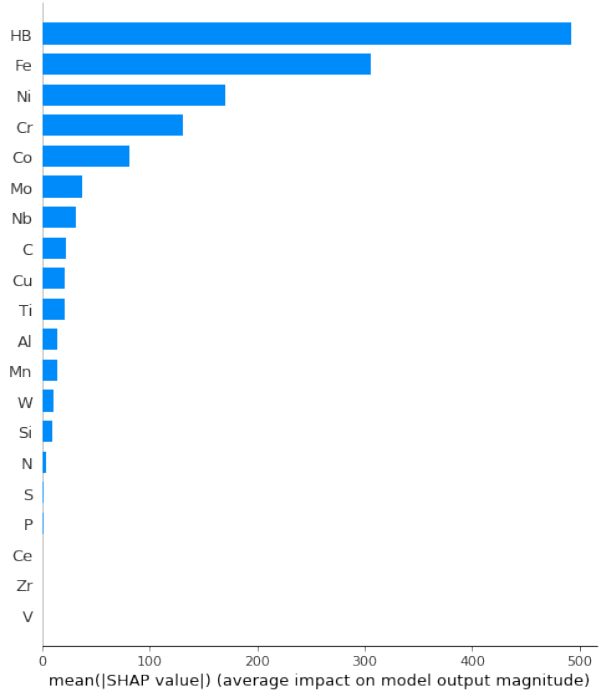

Figure 9. Feature importance calculated by SHAP in NN for Yield Strength.

In the model of FCNN to YS, the three most important feature inputs besides hardness, are Fe, Ni, and Cr. Nickel (Ni) is added to the steel to boost hardenability. Even with the increased strength and hardness, it often enhances the steel's toughness and ductility [21] [22] [23]. These elements are different from the elements found for the RF-YS, Keeping

the importance of the hardness and Fe. C, and Ti are replaced by Ni, and Cr.

The model of FCNN used to determine TS is also developed, trained, and tested. Figures 10a and b, present training set and loss, and Figures 11a and b, present the model inference to the actual, and predicted values, and the scatter plot, in a test set for Tensile Strength. The metrics. and the Loss to evaluate the model are also presented in Table 4

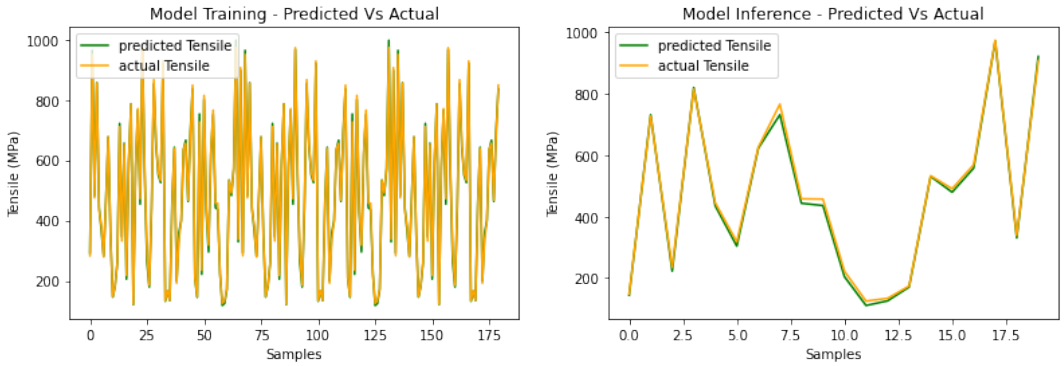

Figure 10. Trainning and Test when calculating Tensile Strength of different steel grades using Neural Network.

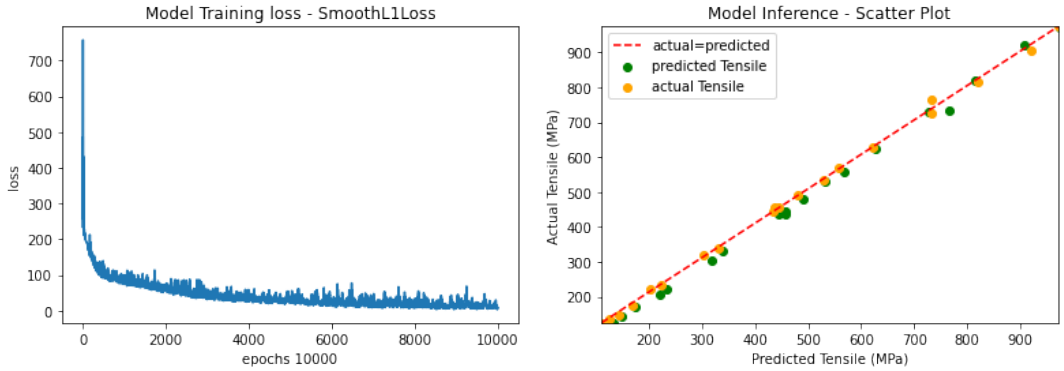

Figure 11. Loss and Inference when calculating Tensile Strength of different steel grades using Neural Network.

Figure 12 shows the feature importance of the SHAP function for NN-TS.

In the model of FCNN to TS, the three most important feature input besides hardness, are Ni, Mo, and Fe. Molybdenum (Mo) affects steel corrosion resistance, as well as hardenability, toughness, and tensile strength. It improves hardenability by lowering the needed quench rate during the heat treatment process, resulting in strong and hard steel. Fe and hardness play a major role in all the models, as it has high values in the data set; however,

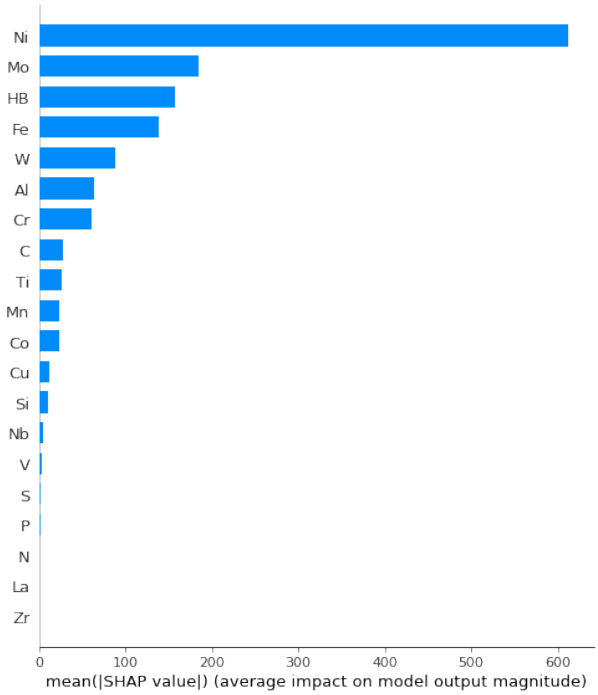

Figure 12. Feature importance calculated by SHAP in NN for Tensile Strength.

each model has different feature importance. In the case of FCNN-TS, they play a lower role than Ni and Mo. If we compare the elements with the RF-TS, the elements of Mo and Ni are replaced by P and Cr with lower importance.

	MAE (MPa)	**MSE (MPa)**	**R2**	**Loss (MPa)**
NN-YS	14.97	23.02	0.99	12.13
NN-TS	10.85	13.22	0.99	8.28

Table 2. Metrics found as to Fully connected Neural Network

The FCNN presents a really good performance, for both YS and TS. With a Low number for MAE, MSE, and R2 close to 1 (one). This model will be also evaluated with lower feature input.

After analyzing the most important elements for determining the mechanical properties, the model of FCNN was tested again with a reduction of feature input in approximately 50% for both RF and NN, that is, twelve features instead of twenty-three. Resultant metrics are shown in Table 3

The features from the SHAP function have the best performance for YS, while the feature

	MAE (MPa)	MSE (MPa)	R2	Loss (MPa)
RF-YS	27.01	54.97	0.93	22.73
RF-TS	19.98	27.04	0.98	31.09
NN-YS	20.72	50.01	0.96	28.71
NN-TS	27.99	40.45	0.97	23.11

Table 3. Metrics found as to FCNN for 50% reduction of features inputs.

of RF has a best performance for TS. In general, with 50% of the original features the model does not increase the error, as the worst R2 was 0.93 and the best was 0.98.

Motivated by the success of the test with 50% features, a reduction of 70% of the feature input was employed, which means that eight features were used as input in the model. Results obtained for each metric are shown in Table 4

	MAE (MPa)	MSE (MPa)	R2	Loss (MPa)
RF-YS	17.20	54.29	0.95	16.70
RF-TS	65.51	116.5	0.79	46.21
NN-YS	37.49	70.71	0.92	37.37
NN-TS	16.6	22.5	0.99	23.61

Table 4. Metrics found as to FCNN for 70% reduction of features inputs.

In this case, the features from RF had the best performance for the YS, and, on the opposite side, the SHAP function has the best performance for TS. In general, with 30% of features the model increases considerably the error in the RF value for TS, increasing the MSE from 27.04 to 116.5. In this instance, the worst case has 0.79 R2, and the best has 0.99 R2.

5 CONCLUSIONS

This article presented two different methods to measure the feature importance: Random Forest and Shapley additive Explanations. Of these two, Random Forest is more commonly used, because it is easily approached, and older than the SHAP method. SHAP is a recent method, published for the first time in 2017 [16]. The method was developed to improve the algorithm of feature selection of Shapley Values. This new technique is used to measure the feature importance in "Black Box" algorithms as a NN. In this work, the Fully neural Network model was employed together with SHAP.

- Four models were developed, two FCNNs to measure the YS and TS, and two RF to measure YS, and TS.

- Besides that, it was present in this article four graphs with variable importance for RF to YS, and TS. And FCNN to YS, and TS.

- The both methods presented great results for the suggested regression problem.

- The different methods present different feature importance.

- With 50% reduction of feature input still presents great results.

- However, when we have a 70% reduction of inputs, the results in NN get worse with values of RF algorithm to tensile, and for yields strength, the Values from NN present worst results.

- Both methods are capable to find great results and analyzing the data; however, in this work, the SHAP method shows a bit better performance than RF values.

- There are some qualitative differences in the results obtained for both models in terms of which chemical elements play a major role in the studied mechanical properties. Hence for future work an extension of the database is concerned, as well as an application of new methods of feature selection trying to explain those differences.

REFERENCES

[1] J. L. Cann, A. De Luca, D. C. Dunand, D. Dye, D. B. Miracle, H. S. Oh, E. A. Olivetti, T. M. Pollock, W. J. Poole, R. Yang, *et al.*, "Sustainability through alloy design: Challenges and opportunities," *Progress in Materials Science*, vol. 117, p. 100722, 2021.

[2] "World steel in figures 2019." World Steel Association, 2019:3–30.

[3] S. S. Sidhu, H. Singh, and M. A.-H. Gepreel, "A review on alloy design, biological response, and strengthening of β-titanium alloys as biomaterials," *Materials Science and Engineering: C*, vol. 121, p. 111661, 2021.

[4] N. Yurchenko, E. Panina, S. Zherebtsov, and N. Stepanov, "Design and characterization of eutectic refractory high entropy alloys," *Materialia*, vol. 16, p. 101057, 2021.

[5] H. Li, A. Wang, T. Liu, P. Chen, A. He, Q. Li, J. Luan, and C.-T. Liu, "Design of fe-based nanocrystalline alloys with superior magnetization and manufacturability," *Materials Today*, vol. 42, pp. 49–56, 2021.

[6] Y. T. Tang, C. Panwisawas, J. N. Ghoussoub, Y. Gong, J. W. Clark, A. A. Németh, D. G. McCartney, and R. C. Reed, "Alloys-by-design: Application to new superalloys for additive manufacturing," *Acta Materialia*, vol. 202, pp. 417–436, 2021.

[7] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning*, vol. 4. Springer, 2006.

[8] E. Alpaydin, *Machine learning*. MIT Press, 2021.

[9] H. Liu and H. Motoda, *Computational methods of feature selection*. CRC Press, 2007.

[10] C. Kingsford and S. L. Salzberg, "What are decision trees?," *Nature biotechnology*, vol. 26, no. 9, pp. 1011–1013, 2008.

[11] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[12] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt, and G. Varoquaux, "API design for machine learning software: experiences from the scikit-learn project," in *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pp. 108–122, 2013.

[13] T. Szandała, "Review and comparison of commonly used activation functions for deep neural networks," in *Bio-inspired neurocomputing*, pp. 203–224, Springer, 2021.

[14] M. Seabra and A. Costa, "Material model calibration using machine learning: a comparative study," *European Journal of Computational Mechanics*, 2022.

[15] R. Rojas, "The backpropagation algorithm," in *Neural networks*, pp. 149–182, Springer, 1996.

[16] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in neural information processing systems*, vol. 30, 2017.

[17] M. T. Ribeiro, S. Singh, and C. Guestrin, "" why should i trust you?" explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 1135–1144, 2016.

[18] S&G, "Steel-grades." https://www.steel-grades.com, 2021 [online].

[19] C. Labrecque and M. Gagne, "Ductile iron: Fifty years of continuous development," *Canadian metallurgical quarterly*, vol. 37, no. 5, pp. 343–378, 1998.

[20] A. Alhussein, M. Risbet, A. Bastien, J.-P. Chobaut, D. Balloy, and J. Favergeon, "Influence of silicon and addition elements on the mechanical behavior of ferritic ductile cast iron," *Materials Science and Engineering: A*, vol. 605, pp. 222–228, 2014.

[21] R. A. Gonzaga and J. F. Carrasquilla, "Influence of an appropriate balance of the alloying elements on microstructure and on mechanical properties of nodular cast iron," *Journal of Materials Processing Technology*, vol. 162, pp. 293–297, 2005.

[22] O. Gvozdeva, A. Shalin, and A. Stepushin, "The correlation among chemical composition, structure and mechanical properties in titanium alloys for the elements with increased dynamic ability," in *IOP Conference Series: Materials Science and Engineering*, vol. 709, p. 022082, IOP Publishing, 2020.

[23] W. Zhuang, H. Zhi, H. Liu, D. Zhang, and D. Shi, "Effect of titanium alloying on the microstructure and properties of high manganese steel," in *E3S Web of Conferences*, vol. 79, p. 01001, EDP Sciences, 2019.

[24] T. Saburi, T. Tatsumi, and S. Nenno, "Effects of heat treatment on mechanical behavior of ti-ni alloys," *Le Journal de Physique Colloques*, vol. 43, no. C4, pp. C4–261, 1982.

[25] B. Wang, R. Xin, G. Huang, and Q. Liu, "Effect of crystal orientation on the mechanical properties and strain hardening behavior of magnesium alloy az31 during uniaxial compression," *Materials Science and Engineering: A*, vol. 534, pp. 588–593, 2012.