XXVII Congresso Nacional de Estudantes de Engenharia Mecânica
08 a 12 de fevereiro de 2021, Curitiba, PR, Brasil

Prediction of yield and tensile strengths for high-alloy steels from chemical composition: a data preprocessing approach

Giovanni Corsetti Silva, corsetti@ufpr.br[1]
Diogo Berta Pitz, diogo.pitz@ufpr.br[1]

[1]Universidade Federal do Paraná, Av. Coronel F. H. dos Santos, 100 – Jardim das Americas, Curitiba, PR, Brasil

Abstract. *This paper presents the prediction of both yield and tensile strengths for high-alloy steels from chemical composition only, that is, the microstructure, grain size, temperature and other variables are unknown. For that purpose, a neural network was designed, where the input features were: i) raw data and; ii) Regression F-test filtered data. For both cases, a Bayesian optimizer was utilized for tuning. The results show that for the yield strength, preprocessing the data through an F-test gives the best results (R-squared equal to 0.84), meanwhile for the tensile strength, the raw data produces the best performance (R-squared equal to 0.85). Depending on the target and dataset size, removing features through an F-test can substantially increase the accuracy, however, in the scenario where important features are removed, the accuracy dwindles. It was also observed that a convenient way to choose the most efficient neural network is selecting the one with the lowest number of neurons, indicating low overfitting.*

Keywords: *High-Alloy Steels. Yield Strength. Tensile Strength. F-Test. Neural Network.*

1. INTRODUCTION

In the past few years artificial intelligence (AI) has been gaining popularity in different areas of science and becoming especially important for engineers and data scientists. Among all the tools in AI, neural network (NN) is undoubtedly one of the most popular, mainly due to its capacity of solving complex and massive problems by learning from experience (Hinton, 1992). Neural networks have the most diverse applications, such as tropical cyclone forecasting (Liu, *et al.*, 2005), stock prediction (Liu and Ma, 2012) and face recognition (Aitkenhead and Mcdonald, 2003). They permit the development of quantitative expressions without compromising the known complexity of the problem (Bhadeshia, 2009).

In the field of materials science, there are difficult problems where the general concepts might be understood but which are not as yet amenable to scientific and mathematical treatment (Bhadeshia, 1999). Because of that, in the past few years neural networks have been widely used in materials science to help solve those tough tasks. As an example, Singh *et al.* (1998) predicted the yield to tensile strength ratio as a function of the carbon and manganese concentrations of hot-rolled steels, and those results are now used for the automation of rolling mills (Alaei, *et al.*, 2016). Yield and tensile strengths are of paramount importance for structural integrity and reliability (Karmazínová and Melcher, 2012), and its accurate prediction involves a huge number of parameters, such as the chemical composition (Calik, *et al.*, 2010; Wang, *et al.*, 2017), grain size (Margolin and Hashimoto, 1981; Rice, 1993), temperature (Wang, *et al.*, 2013), microstructure (Mizuno, *et al.*, 2010) and others. Although an accurate prediction should involve many parameters, as mentioned before, there is a strong, practical need for an estimation of yield and tensile strengths utilizing only the chemical composition, without any additional investigation of the detailed process that should control other features.

Demura *et al.* (2019) predicted the creep rupture time for steels using as inputs only the chemical composition and test conditions, obtaining a reasonable prediction, extremely useful as an estimate, since in the real world there may not be further information available apart from the chemical composition of the steel. Following Demura's idea, in the present work the authors estimate both the yield and tensile strengths for high-alloy steels only from chemical composition applying statistical methods and a neural network.

1.1. Feed-forward neural network model

A feed-forward neural network is the simplest type of artificial neural network devised, where the information moves only forward, from the input nodes through the hidden nodes and finally to the output nodes. Fig. 1 represents a simple feed-forward neural network.

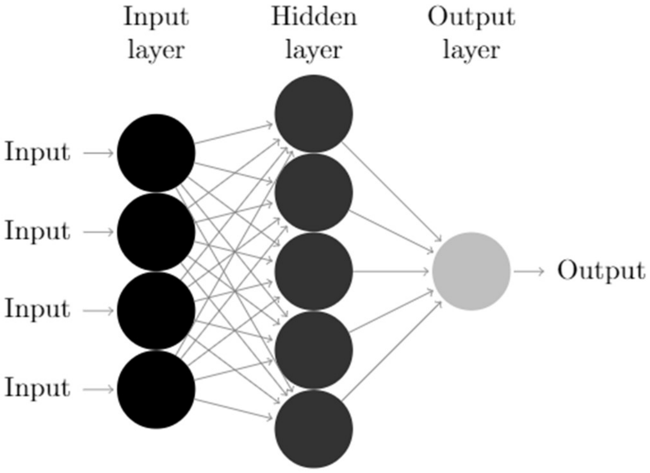

Figure 1. Feed-forward neural network (author, 2020)

The input layer receives correspondent values to its input X_i (i=1,2,…,N), where N is the number of X features, also corresponding to the number of input nodes. The output of the hidden layer, f(z), is computed from the values of the previous layer weighted, as can be seen in Eq. (1), and transformed by a pre-specified activation function (AF), represented in Eq. (2) as follows:

$$z = w_0 + \sum_{i=1}^{n} x_i w_i \tag{1}$$

$$f(z) = AF(z) \tag{2}$$

When the hyperparameters of a neural network are selected, the weights are then tuned through the back-propagation algorithm, that adjusts the 'strength' of the connections between nodes and reduces the error between the desired signal and output of the neural network. To the best of the authors' knowledge, there is no specific method to determine the optimal hyperparameters for a neural network, and most optimization approaches rely on trial and error. Among the available optimization algorithms, the Bayesian optimizer carries out a probability search based on results from prior random runs and converges to the optimal hyperparameters, being widely used. Although the Bayesian algorithm does tweak the hyperparameters, the selected range of hyperparameters that will be analyzed must be specified beforehand, and the wider it is, the longer it will take for the algorithm to converge.

1.2. Data reduction

A drawback regarding a neural network is its heavily dependence on a large quantity of data. More data will always result in better and more reliable results (Dubost, *et al.*, 2019; Xu and Liu, 2019). Unfortunately, gathering large datasets with manually labeled data is both work-intensive and expensive (Wagner, *et al.*, 2013). When increasing the dataset is virtually inviable, data reduction is a good approach to compact the model (Ali, 2004; Fyefe, 1997). For few observations with many features data reduction proves to be extremely powerful, since when the neural element number n of neural networks is larger than the sample size m, the overfitting problem arises because there are more parameters than actual data (more variables than constraints) (Zhang, *et al.*, 2015). The Compression of neural network input data makes it possible to design smaller neural networks than those without data compression (Kuzniar and Zajac, 2015), and currently the most popular method for data dimensionality reduction is based on Principal Component Analysis (PCA) (Tan and Mayrovouniotis, 1995).

Principal Component Analysis (PCA) and variant methods are dimension reduction techniques that rely on orthogonal transformations in order to find a low-dimensional set of axes that compress the data based on the variance of each feature (Battaglino and Koyuncu, 2020; Hotelling, 1933). Generally speaking, PCA, instead of simply identifying the most important variable, combines variables into new constructs made up of different variables and then identifies the most important of the new constructs (Emerson, 2020). A drawback in PCA method is that it cannot be used in design problems, since features tend to be independent from each other.

For small datasets where the features are not correlated, a feasible simplifying alternative is to eliminate features with low impact in the model, performing a bias-variance tradeoff. One way of analysing the relevance of parameters is through an F-test, which compares two regression models based on an analysis of variance (ANOVA). Consider the following unrestricted model, illustrated in Eq. (3) bellow:

$$\beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n = y \tag{3}$$

To test the significance of one or more parameters, the model is fitted without the parameters being analysed, as can be seen in the restricted model bellow, represented by Eq. (4):

$$\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k = y \tag{4}$$

When both models are compared, the null hypothesis that states that the feature has no influence in the model is then tested, according to Eq. (5):

$$H_0\text{:}\ \sum_{i=k}^{n} \beta_i = 0 \tag{5}$$

The number of features in the model is represented by n and k is the number of features considered after applying a restriction on it. In order to evaluate if the restricted model fails to explain the variance in the target value it is necessary to reject the null hypothesis (H_0), and this can be made by utilizing the following test statistic, represented in Eq. (6) in both models:

$$\hat{T} = \frac{(SSR_R - SSR_U)/(n-k)}{SSR_U/(m-n+1)} \tag{6}$$

Where SSR_R and SSR_U represent the sum of squared residuals for the restricted and unrestricted models, respectively, and m denotes the sample size. The terms (n-k) and (m-n+1) are called "degrees of freedom". The test statistic $\hat{T}$ follows an F-distribution, therefore its p-value can be obtained from the F-distribution. If the p-value is below a predetermined threshold (usually 0.05, by Fisher's standards) the null hypothesis can be rejected and the restricted model disregarded. An important observation is that the errors from the two models must be independent and normally distributed. If the assumption of error independence is violated the model is invalid, however even if the errors are nonnormally distributed the F-test remains robust for a dataset that is not too small (Ali and Sharma, 1996).

2. MATERIALS AND METHODS

A dataset containing 312 high-alloy chemical compositions, in percentage (%), and yield and tensile strengths, in MPa, was obtained from (Ward, *et al.*, 2018). The dataset was created by Gareth Conduit of Cambridge University and Intellegens. The features that will be utilized as inputs are the chemical elements carbon (C), Manganese (Mn), Silicon (Si), Chromium (Cr), Nickel (Ni), Molybdenum (Mo), Vanadium (V), Nitrogen (N), Niobium (Nb), Cobalt (Co), Tungsten (W), Aluminium (Al) and Titanium (Ti). The target values of the models will be the yield strength (YS) and tensile strength (TS). The most relevant information regarding the dataset are summarized below, in Tab. 1.

Table 1. Characteristics of the dataset. The input features are carbon (C), Manganese (Mn), Silicon (Si), Chromium (Cr), Nickel (Ni), Molybdenum (Mo), Vanadium (V), Nitrogen (N), niobium (Nb), Cobalt (Co), Tungsten (W), Aluminium (Al) and Titanium (Ti). The target values are Yield Strength (YS) and Tensile Strength (TS) (Author, 2020)

Units: 312								
	C	**Mn**	**Si**	**Cr**	**Ni**	**Mo**	**V**	**N**
Mean	0.096	0.146	0.221	8.044	8.184	2.766	0.184	0.006
Std	0.109	0.397	0.581	5.426	6.337	1.833	0.452	0.018
Min	0.000	0.01	0.010	0.010	0.010	0.020	0.000	0.000
Max	0.430	3.000	4.750	17.500	21.000	9.670	4.320	0.150
	Nb	**Co**	**W**	**Al**	**Ti**	**YS**	**TS**	
Mean	0.035	7.009	0.161	0.239	0.311	1420.998	1641.653	
Std	0.162	6.254	0.920	0.340	0.557	301.894	346.475	
Min	0.000	0.010	0.000	0.010	0.000	1005.900	1019.000	
Max	2.500	20.100	9.180	1.800	2.500	2510.300	2570.000	

The dataset was randomly shuffled and 15% of it was removed - also randomly - in order to split a train set and test set. Both the train and test sets were then normalized based on the parameters of the train set. Two different situations were analysed: 1) the neural network was fed with the raw inputs, 2) The train set was fitted in a linear regression model and an F-test was performed under the null hypothesis that a specific feature is redundant for the model, for each feature individually, and a neural network was fed only with the relevant features.

2.1. F-test in the dataset

Table 2 summarizes the p-values under the null hypothesis where the specific coefficient has no significative influence over a simple linear model, which can be used to rule out some inputs from the neural network and reduce its complexity. The reference is the distribution of $F_{1,251}$ for a 5% asymptotic level by the test $\Psi_{5\%}$. As can be seen, the level of 5% is respected for the mutual acceptance of all the null hypothesis according to the sharper Bonferroni correction for multiple tests of significance (Hochberg, 1988).

Table 2. Hypothesis testing of no dependence of each input feature on the train set in a simple linear regression for $F_{1,251}$ and 5% asymptotic level (Author, 2020)

Restriction	p-value (Yield Strength)	H_0: $\beta_n = 0$	p-value (Tensile Strength)	H_0: β_n=0
C	e-11	Reject	e-20	Reject
Mn	0.981	Fail to Reject	0.533	Fail to Reject
Si	0644	Fail to Reject	0.009	Reject
Cr	0.336	Fail to Reject	0.013	Reject
Ni	0.022	Reject	0.027	Reject
Mo	0.789	Fail to Reject	0.106	Fail to Reject
V	0.082	Fail to Reject	0.105	Fail to Reject
N	0.268	Fail to Reject	e-06	Reject
Nb	0.379	Fail to Reject	0.768	Fail to Reject
Co	e-07	Reject	e-19	Reject
W	0.854	Fail to Reject	0.340	Fail to Reject
Al	e-07	Reject	e-06	Reject
Ti	e-20	Reject	e-27	Reject

Based on the previous analysis, the neural network contains different inputs for the yield strength and tensile strength after performing the regression F-test. For the yield strength, the considered features are C, Ni, Co, Al and Ti, and for the Tensile Strength they will be C, Si, Ni, N, Co, Al, Ti. For further confirmation, an $F_{8,251}$-test and an $F_{5,251}$-test were both performed considering the removal of features which could not be rejected in the $F_{1,251}$ test for each case. The results are summarized in Tab. 3, where it can be seen that removing these features indeed does not violate the 5% asymptotic level, as previously mentioned above.

Table 3. Hypothesis testing of no dependence of the removed features on the train set in a simple linear regression for a 5% asymptotic level. For yield strength features an $F_{8,251}$-test was performed, meanwhile for the tensile strength case an $F_{5,251}$-test was performed (Author, 2020)

Target value	Input features	Removed features	p-value	H_0: $\Sigma\beta_n = 0$
Yield Strength	C, Ni, Co,Al, Ti	Mn, Si, Cr, Mo, V, N, Nb, W	0.585	Fail to Reject
Tensile Strength	C, Si, Cr, Ni, N, Co, Al, Ti	Mn, Mo, V, Nb, W	0.333	Fail to Reject

2.2. Algorithm

As previously mentioned, two different input situations are analysed. These conditions are: i) the raw data; ii) data features that were not ruled out in the 5% asymptotic level of the regression F-test. The data was randomly shuffled and split into train and test sets to avoid selection bias. To find the hyperparameters, a Bayesian optimizer was applied. The top 3 results provided by the optimizer were then statistically tested to select the best possible given combination of hyperparameters of the neural network. A summary of the algorithm is described in Tab. 4.

For small dataset sizes, the weight tuning process needs to see the trainset multiple times, which leads to a large number of epochs, that is, the number of times that the model sees the same data. For few data, it is a good approach to set an early stopping training which can be determined empirically. Pons *et al.* (2019) set its value to 200, meanwhile for the dataset presented in this work the value of 100 seemed to work better.

Table 4. Structure of the algorithm utilized in the present work (Author, 2020)

```
Algorithm
Input: Chemical composition of high-alloy steels
        → i) Raw data features
        → ii) data features not ruled out in the F-test$_s$
Output: Yield Strength (YS) / Tensile Strength (TS)
1: shuffle and randomly remove 15% of the dataset to make a trainset and test set
2: normalize both the trainset and test set based on the train set
3: Find hyperparameters with Bayesian optimizer (50 runs)
        → layers: up to 3
        → neurons per hidden layer: 32 to 512, step=32
        → activation functions: RELU, ELU, SELU
        → epochs: 50 to 1000, step=100
        → early stop: patience=100
        → validation test = 15%
        → loss = MSE
4: top 3 best models are tested and the best one picked up
5: end
```

3. RESULTS AND DISCUSSION

Table 5. Optimal hyperparameters found with Bayesian optimizer for all cases (Author, 2020)

		Yield Strength		Tensile Strength	
Layer	**Hyperparameter**	**F-Test**	**Raw Data**	**F-Test**	**Raw Data**
1st hidden layer	Neurons	128	448	256	64
	Activation	RELU	RELU	RELU	RELU
2nd hidden layer	Neurons	64	160	448	480
	Activation	SELU	RELU	ELU	RELU
3rd hidden layer	Neurons	96	192	192	194
	Activation	RELU	SELU	SELU	RELU
	ΣNeurons	288	800	896	738

The optimal neural network hyperparameters for each case are described above in Tab. 5. Depending on how the input is manipulated, it is notable the divergence in the optimal hyperparameters found with the Bayesian optimizer, indicating that every time the input features are shaped, a new hyperparameter search must be carried out to assure the optimal performance. Regarding the total number of neurons, Arifin *et al.* (2019) state that if this value is too large, the neural network performs poorly, although Fumumoto *et al.* (2017) obtained best performances for problems with greater complexity by increasing the number of neurons. It is reasonable to conclude that, since the current problem involves essentially the same task (prediction of yield and tensile strengths), the settings with the lower number of neurons tend to avoid overfitting better. For the yield strength case, shrinking the input features through an F-test displays the lower number of neurons, meanwhile, for the tensile strength case, feeding the neural network with raw inputs results in the smallest number of neurons, which points out the mentioned cases as the ones that will most likely perform better in the test set.

Figure 2 presents subtle results regarding the stability of each condition. The empirical Cumulative Distribution Function (CDF) describes the likelihood of occurrence of an error being less to or equal to a chosen value. For the yield strength errors, the neural network fed with features shrank by an F-test presents high stability, with its median placed practically on the zero position. On the other hand, the raw input has a strong trend to under evaluate the prediction, which could be useful for safety purposes. For the tensile strength errors, both distributions behave with great similarity for the specified range.

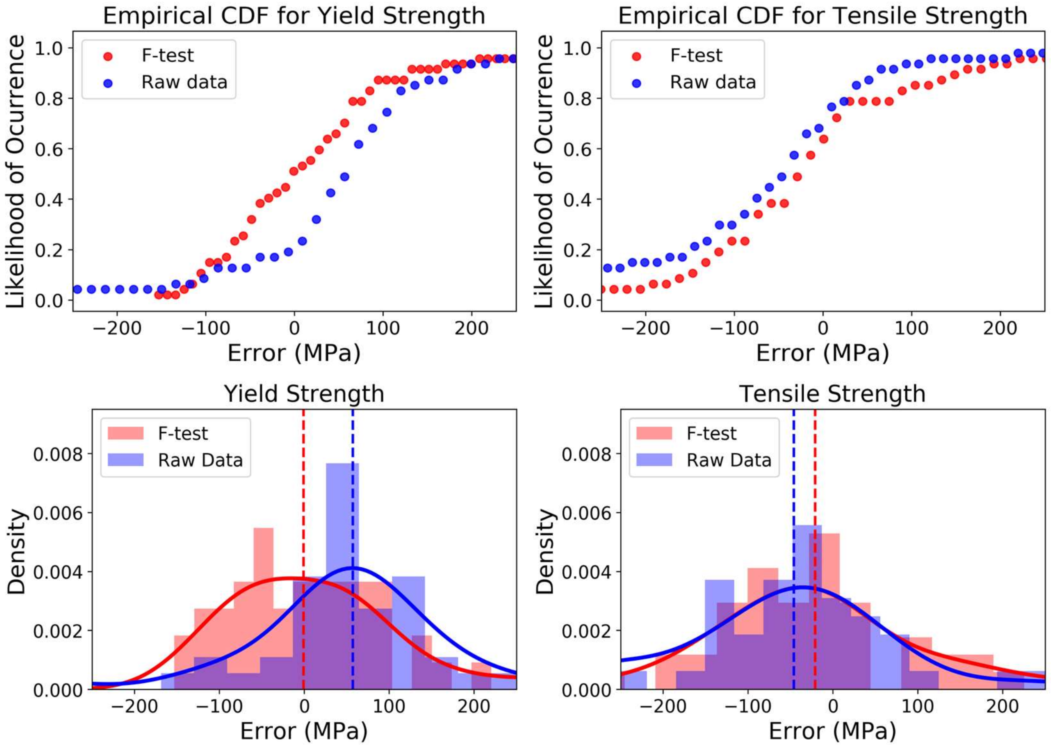

Figure 2: On the top, the empirical Cumulative Distribution Function (CDF) of the errors; on the bottom, the histogram and density plots for the errors, where the dashed lines correspond to the medians for each case (Author, 2020)

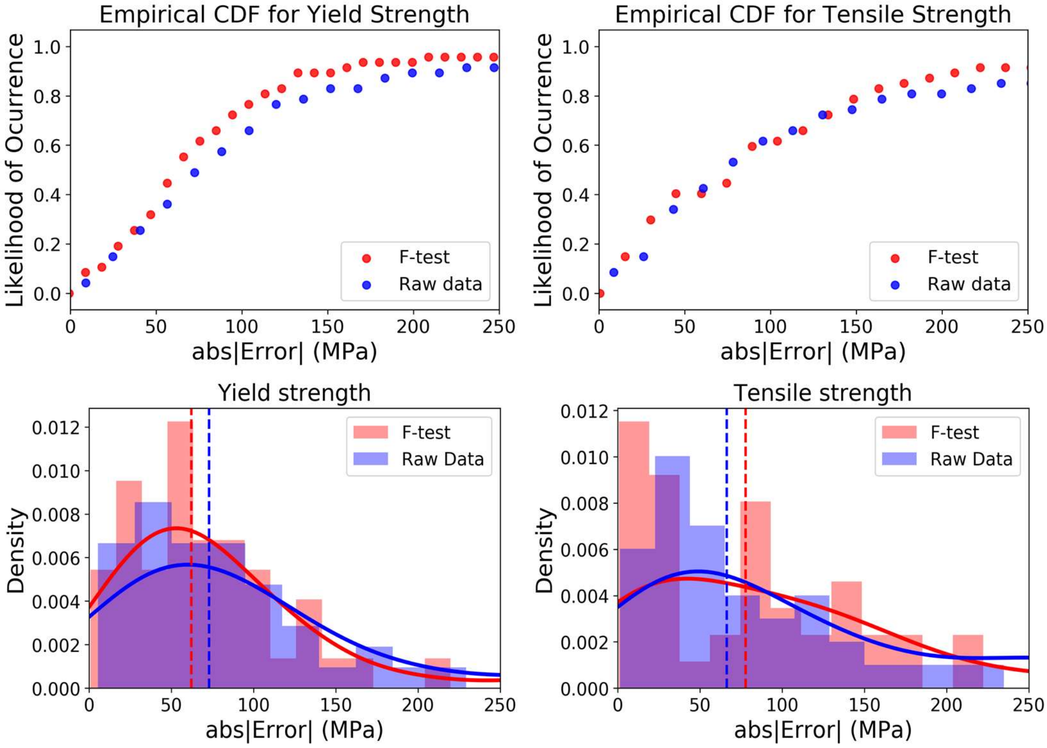

Figure 3: On the top, the empirical Cumulative Distribution Function (CDF) of the absolute value of the errors; on the bottom, the histogram and density plots for the absolute value of the errors, where the dashed lines correspond to the medians for each case (Author, 2020)

The absolute value of the error for both the yield and tensile strengths are presented in Fig. 3. As can be seen for the yield strength model, the F-test case tends to present a slightly lower absolute value of error likelihood, bolstering the previous belief it was going to perform better based on the smaller quantity of neurons. For the tensile strength target, the absolute value of error likelihood for both cases behave with great similarity, therefore a good alternative to analyse which model performs better in practice is by heavily penalizing outliers. The R-squared, described in Fig. 4, heavily penalizes outliers and can be used together with the previous charts to describe the consistency of the models. The yield strength predicted after the F-test filtering presents a higher R-squared value, equal to 0.84, which was already expected from the median and density plot analysis previously mentioned. Regarding the tensile strength prediction, the raw input data model outstands the one filtered with an F-test, presenting an R-squared value equal to 0.85.

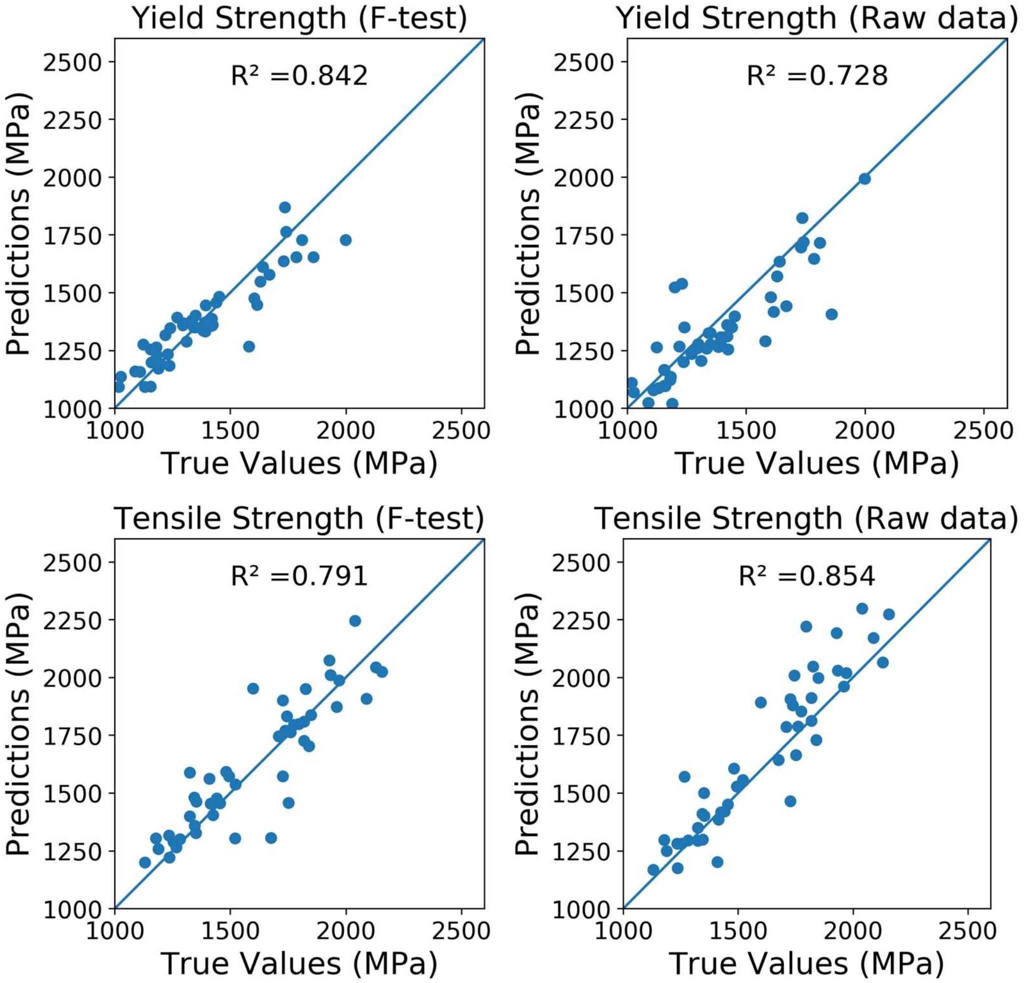

Figure 4: Predicted values plotted against the True Values for all the analysed cases for the yield strength, on the top, and tensile strength, on the bottom, where the blue line splitting the charts represents the perfect predicted condition. On the top of each chart is displayed the R-squared values (Author, 2020)

4. CONCLUSIONS

From the previous analysis, some conclusions can be made:

- Depending on the target and how the features are fed to the Neural network, the optimal hyperparameters can largely diverge;
- For yield strength prediction with few data points, performing an F-test filtering on the input features before feeding them to the Neural network presented the best performance, with an R-squared equal to 0.84;
- If it is desirable to have slightly under evaluated predictions at the expense of some bias for safety purposes during yield strength prediction, the raw data may be used;
- For the tensile strength prediction, the model that received the raw data performed better for having less outliers, with an R-squared of 0.85
- When the target object is maintained the same and transformations are performed in the dataset features, a rough test to check which case will present better results is counting the total number of neurons and picking up the one with the lowest value.

5. REFERENCES

Aitkenhead, M. J. and Mcdonald, A. J. S., 2003. “A neural network face recognition system”. *Engineering Applications of Artificial Intelligence*, v. 16, n. 3, p. 167–176.

Alaei, H., Salimi, M. and Nourani, A., 2016. “Online prediction of work roll thermal expansion in a hot rolling process by a neural network”. *The International Journal of Advanced Manufacturing Technology*, v. 85, n. 5–8, p. 1769–1777.

Ali, A. H., 2004. “Application of Neural network Principal Components to Climate Data”. *Journal of Atmospheric and Oceanic Technology*, v. 21, n. 2, p. 149–158.

Ali, M. M. and Sharma, S. C., 1996. “Robustness to nonnormality of regression F-tests”. *Journal of Econometrics*, v. 71, n. 1–2, p. 175–205.

Arifin, F., Robbani, H., Annisa, T. N. and Ma'arof, N. N. M. I., 2019. “Variations in the Number of Layers and the Number of Neurons in Artificial Neural networks: Case Study of Pattern Recognition”. *Journal of Physics: Conference Series*, v. 1413, p. 012016.

Battaglino, S. and Koyoncu, E., 2020. “A Generalization of Principal Component Analysis”. *International Conference on Acoustics, Speech and Signal Processing - ICASSP 2020*. Barcelona, Spain.

Bhadeshia, H. K. D. H., 1999. “Neural networks in Materials Science”. *ISIJ International*, v. 39, n. 10, p. 966–979.

Bhadeshia, H. K. D. H., 2009. “Neural networks and Information in Materials Science”. *Statistical Analysis and Data Mining*, v. 1, n. 5, p. 296–305

Calik, A., Duzgun, A., Sahin, O. and Ucar, N., 2010. “Effect of Carbon Content on the Mechanical Properties of Medium Carbon Steels”. *Zeitschrift für Naturforschung A*, v. 65, n. 5, p. 468–472.

Demura, M., Sakurai, J., Yamazaki, M. and Inoue, J., 2019. “Machine Learning Prediction of Creep Rupture Time for Steels”. In *International Conference on Computational & Experimental Engineering and Sciences*, v. 22, n. 2, p. 123–123.

Dubost, F., Bortsova, G., Adams, H., Ikram, M. A., Niessen, W. J., Vernooij, M. and de Bruijne, M., 2019. “Hydranet: Data Augmentation for Regression Neural networks”. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, v. 11767, p. 438–446.

Emerson, R. W., 2020. “A Look at Principal Component Analysis”. *Journal of Visual Impairment & Blindness*, v. 114, n. 3, p. 240–241.

Fumumoto, Y., Owaki, S. and Nakamura, M., 2020. “Effect of number of neurons of a neural-network on compensation performance of SPM non-linear waveform distortion”. In *Opto-Electronics and Communications Conference (OECC) and Photonics Global Conference (PGC)*. Singapore.

Fyfe, C., 1997. “A Neural network for PCA and Beyond”. *Neural Processing Letters*, v. 6, n. 1/2, p. 33–41.

Hinton, G. E., 1992. “How Neural networks Learn from Experience”. *Scientific American*, v. 267, n. 3, p. 144–151.

Hochberg, Y., 1988. “A sharper Bonferroni procedure for multiple tests of significance”. *Biometrika*, v. 75, n. 4, p. 800–802.

Hotelling, H., 1933. “Analysis of a complex of statistical variables into principal components”. *Journal of Educational Psychology*, v. 24, n. 6, p. 417–441.

Karmazínová, M. and Melcher, J., 2012. “Influence of Steel Yield Strength Value on Structural Reliability”. *Recent Researches in Environmental and Geological Sciences*.

Kuzniar, K. and Zajac, M., 2015. “Some methods of pre-processing input data for neural networks”. *Computer Assisted Mechanics and Engineering Sciences*.

Liu, J. N. K. and Feng, B., 2020. "A neural network regression model for tropical cyclone forecast". In *International Conference on Machine Learning and Cybernetics*. Guangzhou, China.

Liu, X. and Ma, X., 2012. "Based on BP Neural network Stock Prediction". *Journal of Curriculum and Teaching*, v. 1, n. 1, p. 15.

Margolin, H. and Hashimoto, K., 1981. "A mechanism for the effect of grain size on yield strength". *Scripta Metallurgica*, v. 15, n. 6, p. 679–682.

Mizuno R., Matsuda, H., Funakawa, Y. and Tanaka, Y., 2010. "Influence of Microstructure on Yield Strength of Ferrite–Pearlite Steels". *Tetsu-to-Hagane*, v. 96, n. 6, p. 414–423.

Pons, J., Serra, J. and Serra, X., 2020. "Training Neural Audio Classifiers with Few Data". In *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Brighton, United Kingdom.

Rice, R. W., 1993. "Correlation of Machining-Grain-Size Effects on Tensile Strength with Tensile Strength-Grain-Size Behavior". *Journal of the American Ceramic Society*, v. 76, n. 4, p. 1068–1070.

Singh, S., Bhadeshia, H. K. D. H., Mackay, D. J. C., Carey, H. and Ar, S. A., 1995. "Neural network Analysis of Steel Plate Processing". *Ironmak Steelmak*, v. 25, p. 22.

Tan, S. and Mayrovouniotis, M. L., 1995. "Reducing data dimensionality through optimizing neural network inputs". *AIChE Journal*, v. 41, n. 6, p. 1471–1480.

Wagner, R., Thom, M., Schweiger, R., Palm, G. and Rothermel, A., 2020. "Learning convolutional neural networks from few samples". In *International Joint Conference on Neural networks (IJCNN)*. Dallas, USA.

Wang, W., Liu, B. and Kodur, V., 2013. "Effect of Temperature on Strength and Elastic Modulus of High-Strength Steel". *Journal of Materials in Civil Engineering*, v. 25, n. 2, p. 174–182.

Wang, X. J., Sun, X. J., Song, C., Chen, H., Han, W. and Pan, F., 2017. "Enhancement of yield strength by chromium/nitrogen alloying in high-manganese cryogenic steel". *Materials Science and Engineering: A*, v. 698, p. 110–116.

Ward, L., Dunn, A., Faghaninia, A., Zimmermann, N., Bajaj, S., Wang, Q., Montoya, J. H., Chen J., Bystrom, K., Dylla, M., Chard, K., Asta, M., Persson, K. A., Snyder, G. F., Foster, I. and Jain, A., 2018. "Matminer: An open source toolkit for materials data mining". *Computational Materials Science*, v. 152, p. 60–69.

Xu, Q. and Liu, A., 2020. "EDA based Deep Neural network Parameter Optimization". In *Proceedings of the 3rd International Conference on Computer Science and Application Engineering - CSAE 2019*. Sanya, China.

Zhang, Y., Li, Y., Sun, J. and Ji, J., 2015. "Estimates on compressed neural networks regression". *Neural networks*, v. 63, p. 10–17.

6. RESPONSIBILITY FOR INFORMATION

The authors are solely responsible for the information included in this work.