International
Business Management
MEDWELL PUBLICATIONS
SINCE 2000

Data Driven: The Decision-Making Improved by the Strength of Data

Alisson Paulo De Oliveira and Hugo Ferreira Tadeu Braga
Innovation Center, Fundação Dom Cabral, Avenida Princesa Diana, 760, Alphaville, Lagoa dos Ingleses, Brazil

Key words: Data-driven, prediction models, analytics, big-data, artificial intelligence, steelmaking

Corresponding Author:
Alisson Paulo De Oliveira
Innovation Center, Fundação Dom Cabral, Avenida Princesa Diana, 760, Alphaville, Lagoa dos Ingleses, Brazil

Page No.: 203-211
Volume: 15, Issue 05, 2021
ISSN: 1993-5250
International Business Management

Abstract: Data is considered a primary resource for innovation. The existence of a large amount of available data as well as technological tools capable of explore them, allows companies to extract information that can be used to create and implement new ideas and new projects. To this end, the details regarding the care that organizations should have with data are explored. The difficulties regarding the adoption of data-driven approach and some measures to implement this type of decision-making approach are discussed. A real example of prediction model for decision making that is based on industrial data is also discussed. This example shows the difficulties in the preparation of data for the development of these models which confirms that most of the time spent in the construction of predictive models it is due to this step. The use of the data-driven approach allows organizations to obtain superior results in their processes, thus becoming a tremendous competitive advantage and a special strategic factor in a highly competitive market, regardless of the field of activity.

INTRODUCTION

Decision making refers to the attempt to determine which natural state prevails in a system in such a way that one can choose the action that yields the highest value when that state is realized. If that natural state which yields the highest value can be determined with certainty, the decision maker has the best information and his decision process is reduced to a mere optimization problem[1]. Theoretically, improvements in technologies that collect or analyze data or digitization can reduce error in information by decreasing the level of aggregation that makes it difficult to distinguish between possible states or eliminate noise[2]. This study seeks to explore how is the process of decision making based on data. It explores related technologies: Big-Data, the data itself, prediction models, cultural aspects of decision making and how an organization can be considered as a truly data-driven organization. The aspects related to the development of a prediction model based on Artificial Intelligence are also discussed.

Big-data: Currently, digitization allows organizations the possibility to collect and analyze large amounts of data easily and quickly. However, the so-called Big-Data must be strategically managed to optimize the use of analysis in business management and to overcome the risk of transforming the advantages offered by the Information and Communication Technology (ICT) tools into threats[3]. If the exploration of new information is a basis

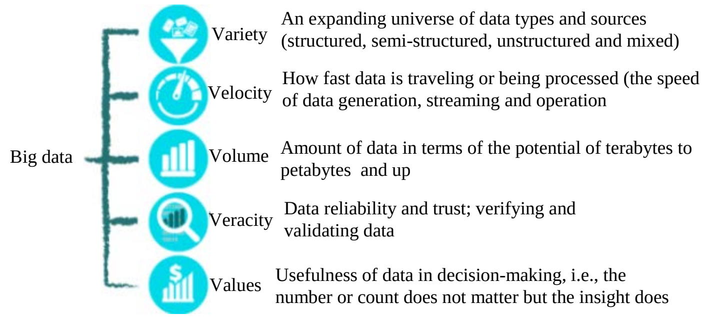

Fig. 1: Big data Fundamentals: the multi-V model[4]

for organizational learning, Big-Data presents an enormous opportunity for companies to learn and consequently, improve their performance[5]. Data is considered a primary resource for innovation[6]. The existence of a large amount of available data as well as technological tools capable of exploiting them, allows companies to extract new information that can be used to create and implement new ideas[7]. In a nutshell, organizational learning through technology like Big-Data can be considered a mixture of data capture, deduction from that data and induction to identify possible patterns which are the basis for possible corrective actions[8]. The fundamentals of Big-Data are explained according to Fig. 1, Multi-V Model[4].

As Big-Data becomes more and more present in the business world, the practical aspects of development and implementing new Data-Driven Business Models (DDBM's) become increasingly important. Therefore, the survival and maintenance of companies competitiveness ends up being dependent on how they intend to use this large amount of data. So, the big challenges are[9]:

- Data extraction
- Refining the data
- Ensure that they are used in the most effective way

Many companies have failed to improve their innovation performance through Big-Data[10] and other companies are still unsure whether Big-Data is really associated with the results obtained[11], although it is considered that companies that use Big-Data in their processes may have a greater opportunity to substantially improve their operational efficiency and revenue compared to their competitors[12]. A recent report shows that in 2016, 48% of companies invested in Big-Data. However, in the same year, the number of companies that stopped using Big-Data was reduced by 6.1%.

Development of a Data-Driven (Empirical) Prediction Model: Predictions model are the base of a truly data-driven organization. Within the field of mathematical modeling there are two distinct branches: the modeling of first principles (or mechanistic) and the empirical modeling. The first branch models are based on a series of equations that examine individual parts of a system[14]. These models are dependent on the understanding of a system, so that, the absence of data is compensated and, therefore have a greater potential for extrapolation in relation to empirical models[15]. The models from the second branch, the empirical models on the other hand, are mathematical equations derived from the analysis of data from a system to be evaluated. However, these models require less knowledge of the system[16]. These empirical models are based on the hypothesis that the available data have sufficient granularity and/or quantity to have the system defined. If this system is defined within the data set, the empirical models derived from these analyzes are valuable tools to characterize the input and output relationships of the system, especially, when there is limited knowledge of the engineering domain, so that complex systems can be characterized. Empirical models are effective, for example, in modeling bioelectro chemical systems which would otherwise require detailed knowledge of complex interactions between the various principles involved: Physical, chemical and electrochemical[17].

These empirical models are already well used in production processes due to the amount of data available[18]. Due to huge advances in the areas of computational intelligence and machine learning, there has been a significant increase in the capacity of empirical models and these new approaches are inserted in the field of data-driven modeling[16]. Computational intelligence seeks inspiration in nature to solve problems[19] or algorithms that replicate human behavior in solving problems such as Artificial Neural Networks[20, 21],

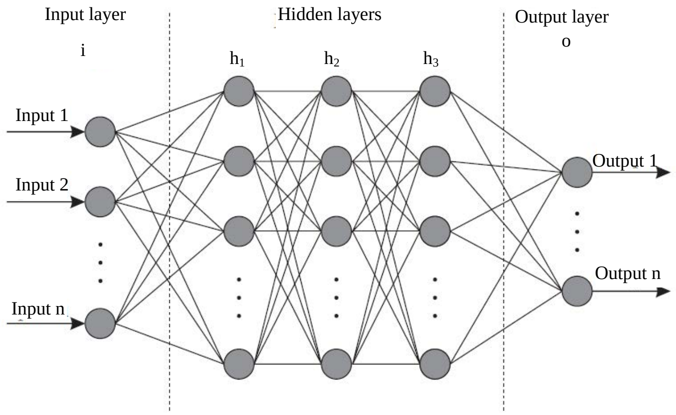

Fig. 2: Artificial Neural Network architecture[22]

(Fig. 2). Data-Drive Models (DDM's) are capable such as Artificial Neural Networks, even without prior knowledge of the system to find relationships between system state variables (Outputs as a function of inputs)[23]. These data-based models train an algorithm (Artificial Neural Networks, Regression) to determine the relationships of a system[20]. Data-based algorithms have been around for a long time[24] but its industrial use was limited by several factors such as restrictions due to lack of data, storage of data in unusable formats and lack of computational processing[25].

It was reported in 2015 that the amount of data generated by organizations, globally was around 1000 exabytes and that just 10 years later this volume would be 10 times greater[26]. The new Cyber-Physical Systems, Smart Factories and Internet of Things applied in industry has caused a huge increase in the volume of data generated[27]. Due to computational costs for modeling large volumes of data, data-driven models were not always considered suitable for modeling. However, technological evolution (Cloud computing, for example) has allowed organizations to access the tools needed to model large data sets. Thus, DDM's have become predominant in organizations aiming at modeling and monitoring industrial processes[25].

Business models: The effective use of data is not only aimed at competitiveness but also the very survival of organizations. Many companies are developing new business models designed specifically to create additional business value by extracting, refining and ultimately capitalizing on the data. Such innovation is notoriously difficult, especially for traditional companies with an ingrained culture. The main motivator, however for conventional companies to become data-driven is the increase in competitive advantage associated with the effective use of Big-Data. Such companies demonstrate an increase in production and productivity of around 5-6% when compared to similar organizations which do not use the data. In the banking sector, 71% of companies report that using Big-Data provides them with competitive advantages. Organizations that fail to make intensive use of data run the risk of losing competitive advantage, even having the possibility of losing market share and revenue[9]. Big-Data can be a source of competitive advantage and a catalyst for successful business models. The three main characteristics of Big-Data: Volume, speed and variety can be considered as sources of competitive advantage in countless new business models[28]. Data-Drive Business Model Innovation (DDBMI) can be defined as the use of information networks and data analysis (Big-Data) from sources that can be internal or external, aiming at the promotion of innovations besides the creation of new business models that aim at monetization through data and even knowledge obtained from external sources. Internal information networks may include data collected from activities inherent to a company's core business model or that data may come from external information networks, such as public sources, Datasets and social media, among other possibilities[29].

Data-driven organizations: In times of Big-Data Analysis, leadership must act as an agent of change within associations[30], constantly dealing with challenges that involve understanding the benefits and availability of data, development of analytical skills and data integration in organizational culture[31]. Leading means improving productivity in associations[32] as well as positively connecting people[33]. Leaders must develop an analytical mindset to transform associations into a decision-making environment that uses data in a very local way[34]. The organizations that are truly data-driven have the following characteristics[35]:

Table 1: Hypothetical issues addressed by the analysis; (d) is a valuable analysis but only (e) and (f) are data driven only if the information is used

Variables	Past	Present	Future
Information	(a) What happened? reporting	(b) What is happening now? alerts	(c) What will happen? extrapolation
Insight	(d) How and why did it happen? modeling, experimental design	(e) What is the next best action? recommendation	(f) What is the best/worst that can happen? prediction, optimization, simulation

Such organizations may be testing on an ongoing basis. These tests may include tests with users where real consumers or users give feedback on new attributes or products.

A truly data-driven organization has a mindset of continuous improvement. They frequently optimize its main processes. And this occurs from the realization of careful analyzes as well as the construction of mathematical or statistical models and the use of these for simulations

Such organizations may be involved in predictive modeling. But even more important, it is the use of model errors as well as other lessons learned in improving the predictive capacity of these models.

A data-driven organization will certainly guide its decisions using a set of weighted variables. The data for each set of variables that are of interest must be collected and the weights between them must be determined to allow the generation of a leadership decision that is reliable.

A truly data-driven organization will have at least one of these characteristics, looking to the future where the data is first-class citizens. An organization that has high quality data in addition to the qualified personnel to analyze it cannot yet be considered as truly data driven. If there is no interest from people in knowing the analyzes and if the decisions of the decision makers are not influenced by these analyzes but by opinions and instinct, it cannot be said that these organizations are data driven. Being data driven means that an organization will use that data as critical evidence to help inform and influence its business strategy. In this way, the culture of this organization will be based on evidence where the data are considered as reliable and the analysis process is highly relevant, informative and used to determine the next steps of the decision-making process[35]. A useful structure for understanding the analysis and its relationship with data-driven is shown in Table 1[35].

This table shows some differences between information and insight, for the past, the present and the future. Report (A) and alert (B) do not use data. They only show that something unusual has happened in the past or at that very moment. The reason is not explained and there is no recommendation on how to avoid the situation. (C) is where eventual extrapolations may not allow precision. (D) is near Data-Driven since it uses prediction models and Design of Experiments (DOE). (E) and (F) represent what Data-Driven really is but just with the use of information because it is the basis for the phenomenon understanding and only if this understanding allows the formulation of an action plan or recommendations to solve the analyzed problem[35].

Decision-making: Decision making using data is a real ideology that considers data to be strategic resources, rather than based on intuition and experience. Such an approach requires the active role of leadership in promoting a culture that always seeks innovation and emphasizes care for data management at each stage of decision making[3]. The following are the factors that hinder the adequate use of data in the decision-making process, thus encouraging the use of experience and instinct[35]:

Factors related to the data: The data must be timely, relevant and reliable. Otherwise, decision makers have limited options. They can postpone the decision, get more data or they can move on and make the decision anyway with whatever data and tools they have which can usually be just experience. Possible problems with the data are:

- Bad data quality and lack of confidence in these
- Volume no possible to be processed
- Sieve the noise signal due to the enormous amount of data

Factors related to culture: Organizational culture is a strong factor and one that evidently influences the way decision making is carried out:

- Valuation of intuition
- Lack of data literacy
- Lack of responsibility

Cognitive barriers: The human being is imperfect and because of that humans are far from being a perfect decision maker. Past experiences and irrelevant details can influence decisions and we often do not approach problems in the most objective way. This whole set can mean completely wrong and possibly not logical decisions. All of this is known as cognitive bias. The reasons why humans cannot always trust our instincts are:

- Inconsistency
- Memories of non-existent events
- The reality is that we are not as good as we think we are

- Humans like bad data
- Humans tie themselves into irrelevant data
- Like any living thing, humans get tired and hungr
- Humans seek to survive
- Confirmation bias
- Bias to always seek the most recent experience
- Bias of friendship or friend or enemy

There are situations where intuition is valued and necessary. Limiting situations such as a burning building at risk of collapse and the evacuation of the fire brigade due to the training and innate sense of danger of a team member. Or even the performance of health professionals who seem to "predict" the future health condition of their patients. However, this type of intuition only occurs in certain environments where the information (Tips, signals) is quite reliable and consistent in addition to the fact that it requires a lot of training and experience from the professional. It takes time to develop and continuous practice often makes perfect[35].As can be concluded, decision-making is a difficult activity as human beings are influenced by several factors (cognitive prejudices, data, culture, etc.). People's ego and their prejudices can influence negatively. However, intuition must be part of an established data-based decision-making process. The relevant data available should always be used and intuition should not be relied on blindly as it can be flawed which would lead to bad decisions. In case of decision making that is contrary to what the data suggest, it should be done with transparency and for the right reasons. On the next section it will be discussed the development of a prediction model, based on an industrial dataset to be used as a Research and Development (R&D) tool for steelmaking companies.

MATERIALS AND METHODS

This study will discuss some steps related to the dataset before the development of a real prediction Artificial Intelligence (AI) Model for steelmaking companies. Such a model allows to know in advance the quality results (mechanical properties) of the steel beam (I-Beam, according to the Fig. 3) used on the construction of buildings, bridges, wind towers for example. In addition, it allows decision makers to adjust rolling processes in view of the initial chemical composition obtained at the melting-shop. Thus, it is a particularly useful tool for quality checking and R&D of new products.

Choice of a structural steel beam for modeling mechanical properties: The mechanical properties of a material are related to its behavior when subjected to external load as well as its ability to resist these efforts without collapse or deforming in an uncontrolled way. For

Fig. 3: Steel I-Beam[36]

this purpose, the Yield Strength (YS), the Tensile Strength (TS) and the Elongation (E) are defined. Among dozens of possibilities of structural beams, the one with the designation W200X46.1 [height (d) X kg/m] was chosen because it presents some interesting particularities, discussed below:

- It has sampling in the flange which implies less variability in the results of the mechanical properties tests
- It is a beam with a reasonable number of tests carried out in various technical standards (ASTM A572-50, ASTM A992, EN 10025-2). Thus, there is a dataset with a larger number of data, ideal for training the AI model

The available dataset: The dataset which was used for training and validating the Artificial Neural Network (ANN), contained data from the production process. Data that allowed the identification of the I-Beams and the type of steel (ASTM A572-50, ASTM A992, ASTM AH36, etc.) was also used. From this dataset, records with the following characteristics were selected:

- Various standards (all related to High Strength and Low Alloy steels)
- Numerous standardized chemical compositions
- Total occurrences equal to 461 (This number refers to the total of samples taken from 207 different steel heats, submitted to tensile tests)

After analyzing the mass of data recovered, it was observed that 11 of the available chemical elements did not present all the values of results indicated in all records. These elements were discarded as input for ANN. They were Ni, Co, Ca, Ti, B, W, Zr, As, Sb, Te and Pb.

Statistical analysis of dataset variables: The dataset obtained from the considerations made in the previous steps has a high number of variables. However, it is expected that a large part of them will be strongly correlated or have a minor influence on the mechanical

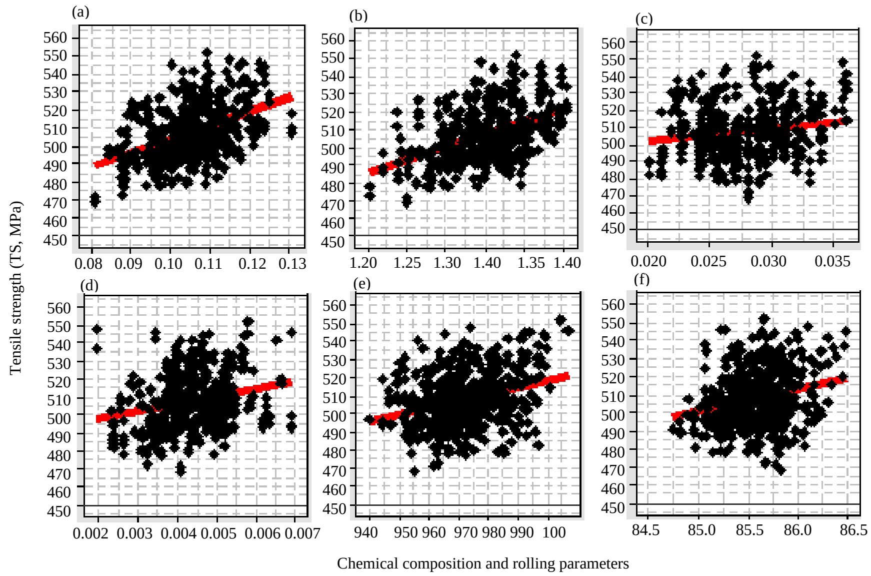

Fig. 4: Tensile strength vs. chemical compositions and process parameters for W200X46.1 steel beam (a) TS, Mpa°Carbon content (b) TS, Mpa°Magnese content % (c) TS, Mpa°Niabium content % (d) TS, Mpa°Nitrogen content% (e) TS, Mpa°final temperature, Flange°C (f) TS, Mpa°Reduction, Flange % Scatterplot of TS, Mpa vs. chemical composition and rolling parameters (b) Chemical composition and rolling parameters

properties of the finished product. For this last situation it is expected that they do not need to be included in the prediction model. To analyze the relationship between the various input variables with the mechanical properties, a statistical analysis of the data obtained was carried out to determine which variables would be used in the training of the ANN. The following steps were carried out on the analyses:

Scatter analysis of the data

- Correlation Analysis between the various input variables (Chemical Composition, Temperatures, Rolling Reductions) and the outputs (YS, TS and E)
- Determination of average, minimum, maximum and standard deviations of the data (Input and output)
- Histograms to check the variability

Data treatment: Statistical analysis of the variables involved was carried out through the MINITAB Statistical Software. The data used were only those that were included within the range +/- 3 Standard Deviations to decrease the total variability of the Dataset.

Elimination of outliers: Elimination of data that was not considered to be representative of the process. In the case of YS and TS, the maximum difference of 20 Mpa was used as an acceptance criterion within the same production order (same steel heat, rolled in the same batch). Events with differences <20 Mpa were excluded.

The techniques mentioned above were used to eliminate the presence of discrepant data, measurement errors in short, noises that could compromise the reliability of the Dataset. Once the mass of data for ANN development was defined, MINITAB was used to graphically analyze the relationship of the output variables (TS, Tensile Strength) with the input variables. The purpose of this procedure was to verify if the impact of the variation of the input data on the Tensile Strength was metallurgically correct. Figure 4 illustrates the dependence of the TS in relation to some of the available process variables. After the steps 1-6 the chosen variables were: C, Mn, Si, S, Cr, Nb, N; Final Temperature and Reduction.

RESULTS AND DISCUSSION

Summary of dataset before and after statistical treatment: After applying the steps 5 and 6 on the original dataset, 444 individual data remained on the new dataset. Table 2 summarizes the percentual change on the standard deviation after performing the steps 5 and 6.

From Table 2, for two of the variables the standard deviation of the final dataset was smaller if compared with the original dataset. For two of the variables the standard deviation increased and for three other variables there was not any change. An additional reduction of the standard variation could be possible but the learning of the ANN could be compromised because of the lack of

Table 2: Statistical summary-original and final dataset

Description	Original dataset SD	Final dataset SD	Change (%)
TS (MPa)	15.58	15.53	-0.32
Carbon (%)	0.0089	0.0090	+1.12
Manganese (%)	0.0508	0.0510	+0.39
Niobium (%)	0.0035	0.0035	+0.00
N2 (%)	0.0009	0.0009	+0.00
Final temperature (°C)	12.45	12.37	-0.64
Reduction (%)	0.31	0.31	+0.00

Table 3: Simulations for artificial neural network with 6 neurons in the hidden layer, tensile strength: Performance parameters

Sim.	Training SSE	Min. error, (%)	Max. error, (%)	Avg error, (%)	R^2, TS (%)
1	3.34946	0.01	7.28	1.52	72.5
2	3.22993	0.03	6.54	1.48	78.2
3	3.42404	0.04	5.17	1.49	77.7
4	3.50430	0.00	6.01	1.47	76.7
5	3.44298	0.02	6.13	1.54	74.4
6	3.27752	0.02	5.28	1.60	77.9
7	3.27218	0.03	6.08	1.48	74.9
8	3.61479	0.01	5.54	1.69	72.3
9	3.41804	0.01	4.86	1.56	73.7
10	2.94769	0.01	6.80	1.83	72.4

Table 4: Summary of characteristics of artificial neural networks

Characteristic	Criteria	MATLAB
Partition of data set	Training set = 75%,	
Validation set = 25%.	RANPERM	
Net weigh initialization	-	INITNW
Net learning ratio	-	TRAINGDX
Transfer function	-	TANSIG
Convergence criteria	-	$SSE = \frac{1}{N}\sum_{p=1}^{N}\sum_{i=1}^{M}\left(t_{pi}-O_{pi}\right)^2$
Minimum error aimed	0.001	-
Number of training cycle	700	-
Training mode	Batch	-
Number of hidden layers	1	-
Size of hidden layer, TS Model	6	-
Net training mode	-	TRAINBR

data. A three-layer ANN architecture was used (One input layer, one hidden layer and one output layer). The optimal number of neurons in the hidden layer was defined using the trial-and-error method and configurations with at least 4 neurons were tested. The variation observed in the training of the networks created difficulties to define the optimal number of neurons in the hidden layer. Thus, the procedure started with 10 ANN simulations performed for each number of neurons using MATLAB. Table 3 show just an example for the results of the simulaton of the ANN for the TS, number of neurons equal to 6 in the hidden layer.

The Analysis of Variance (ANOVA) was performed through MINITAB Statistical Software for the results obtained for each number of neurons in the hidden layer, according the performance parameter (Errors and correlations). From the results obtained, the ideal network configuration was defined. In the ANOVA for the

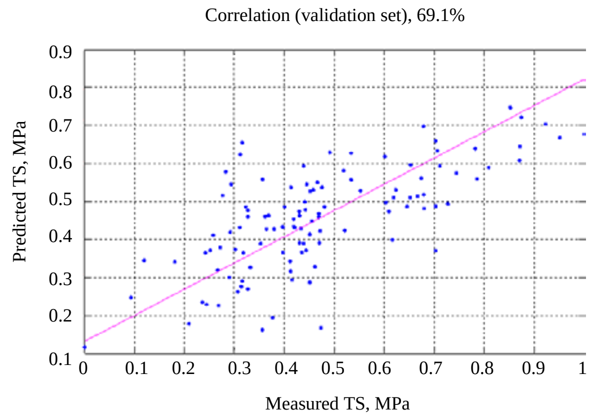

Fig. 5: Tensile strength, TS: predicted vs. measured

performance parameters (minimum, average and maximum error and correlation between TS's), a p-value higher than the significance level ($\alpha = 0.05$) was obtained for the minimum and the maximum error obtained. This result shows that there are no significant differences for each tested configuration. For the average error, a p-value equal to 0.032 was obtained, less than α, indicating significant differences between the configurations. In the case of the correlation, the p-value was also above theα, equal to 0.095. For the average error, a lower value is observed for the configuration with 6 neurons with the standard deviation close to the maximum value obtained among all configurations. As already described, it is possible to state that the average error value varies significantly by using 4 to 8 neurons as a function of the p-value obtained. In the case of the correlation between TS's, a higher average correlation and a lower standard deviation were obtained for the configuration of 6 neurons. With these results, the configuration with 6 neurons was selected for the modeling of the TS. Table 4 shows the summary of the characteristics of the Artificial Neural Network used in this work for TS.

For the TS model (6 neurons on the hidden layer) the correlation between the measured TS value and the predicted TS value was equal to 0.75; the average error was equal to 1.57%, the minimum error was equal to 0,02% while the maximum error was equal to 5.97% (Training set). Another result evaluated was the correlation between the measured and calculated TS values which is shown graphically in Fig. 5. The data refer to the 111 samples used in the validation of the network. In this case, a linear correlation (ρ) equal to 69.1% was obtained.

The final model could be considered a reliable tool to decision making of the steelmaking process once the performance parameters were considered adequate and it was metallurgically sound. The influence of each variable was according to what is expected. A similar procedure was performed for Yield Strength and Elongation with similar results. So, it can be used to calculations of scenarios and decision-making tool from the melting shop to the rolling mill in steelmaking companies. The procedure used to improve the quality of the dataset proved to be adequate and better results could be achieved if there was not a limit to data quantity to ANN training. A larger dataset could offer better results. The use of data proved to be de significative importance once the use of prediction models is a pre-requisite, at least (D) on table 1, to an organization be considered as Data-Driven with better results if compared to non-data-driven organizations.

CONCLUSION

This study discussed the theoretical background related to Data-Drive Decision Making, its correlation with the increasing availability of data and how organizations that use this strategy can deliver better results if compared to organizations that do not uses data. The details necessary for an organization to be in fact data-driven were discussed. It also explores the characteristics of organizations that are truly data-driven, as well as what factors can influence decision-making. Factors that would prevent organizations from being data driven. An example of Data-Driven prediction model was discussed in detail: The development of a real predictive model that enables the decision-taking in a steel industry with minimum error and metallurgically accurate. This last case is discussed with a focus on the noise reduction of the dataset used and the methodology used to adjust the Artificial Neural Network. This article is expected to contribute to the growth of the technical knowledge of its readers.

This is study discusses what is behind a truly data-driven organization, the value of the data, the used technologies, notably the Big-Data and the Prediction Models, how the business models can benefit from the use of data. The data-driven organizations are characterized in details and the decision making process, related to the use of data is also discussed with a focus on the factors that desencourage the decision makers to use data. A real development of a prediction model, for decision making in a steelmaking organization, detailing the optimization of the dataset is discussed and the results from the prediction model developed. The model developed allows the decision makers to have a deep insight into the steelmaking process of a steel I-Beam and actions aiming to know in advance the mechanical properties according to the process parameters are possible, allowing prediction, optimization and simulation.

REFERENCES

01. Blackwell, D., 1953. Equivalent comparisons of experiments. Anal. Math. Stat., 24: 265-272.
02. Brynjolfsson, E. L.M. Hitt and H.H. Kim, 2011. Strength in numbers: How does data-driven decisionmaking affect firm performance?. Econ. Comput. Sci. O&M: Decision-Making Organizations eJournal, Vol. 1, 10.2139/ssrn.1819486
03. Troisi, O., G. Maione, M. Grimaldi and F. Loia, 2020. Growth hacking: Insights on data-driven decision-making from three firms. Ind. Marketing Manage., 90: 538-557.
04. Agrawal, V., 2019. Big data in a nutshell. Software & Support Media GmbH, London, UK.
05. Jones, M., 2019. What we talk about when we talk about (big) data. J. Strategic Inf. Syst., 28: 3-16.
06. Ghasemaghaei, M., 2018. Improving organizational performance through the use of big data. J. Comput. Inf. Syst., 60: 395-408.
07. Ghasemaghaei, M. and G. Calic, 2019. Can big data improve firm decision quality? The role of data quality and data diagnosticity. Decision Support Syst., 120: 38-49.
08. Calvard, T.S., 2016. Big data, organizational learning and sensemaking: Theorizing interpretive challenges under conditions of dynamic complexity. Manage. Learn., 47: 65-82.
09. Brownlow, J., M. Zaki, A. Neely and F. Urmetzer, 2015. Data and analytics-data-driven business models: A blueprint for innovation. University of Cambridge, Cambridge, England.
10. Johnson, J.S., S.B. Friend and H.S. Lee, 2017. Big data facilitation, utilization and monetization: Exploring the 3Vs in a new product development process. J. Prod. Innovation Manage., 34: 640-658.
11. Ghasemaghaei, M., K. Hassanein and O. Turel, 2017. Increasing firm agility through the use of data analytics: The role of fit. Decis. Support Syst., 101: 95-105.
12. Marshall, A., S. Mueck and R. Shockley, 2015. How leading organizations use big data and analytics to innovate. Strategy Leadersh., 43: 32-39.
13. Anonymous, 2016. Gartner survey reveals investment in big data is up but fewer organizations plan to invest. Gartner, Inc., Stamford, Connecticut.
14. Schichl, H., 2004. Models and the History of Modeling. In: Modeling Languages in Mathematical Optimization, Josef, K. (Ed.)., Springer, Boston, USA., pp: 25-36.
15. Mathews, P.G., 2004. DOE Language and Concepts. In: Design of Experiments with MINITAB, Mathews, P.G. (Ed.)., ASQ Quality Press, Milwaukee, Wisconsin, pp: 93-142.
16. Solomatine, D., L. See and R. Abrahart, 2009. Data-Driven Modelling: Concepts, Approaches and Experiences. In: Practical Hydroinformatics, Abrahart, R.J., L.M. See and D.P. Solomatine (Eds.)., Springer, Berlin, Germany, pp: 17-30.

17. Luo, S., H. Sun, Q. Ping, R. Jin and Z. He, 2016. A review of modeling bioelectrochemical systems: Engineering and statistical aspects. Energies, Vol. 9, No. 2. 10.3390/en9020111.
18. Rasmuson, A., B. Andersson, L. Olsson, R. Andersson, L. Olisson and R. Andersson, 2014. Empirical Model Building. In: Mathematical Modeling in Chemical Engineering, Rasmuson, A., B. Andersson, L. Olsson, R. Andersson, L. Olisson and R. Andersson (Eds.)., Cambridge University Press, Cambridge, pp: 40-52.
19. Saka, M.P., E. Dogan and I. Aydogdu, 2013. Analysis of swarm intelligence-based algo-rithms for constrained optimization. Swarm Intell. Bio-Inspired Comput., 1: 25-48.
20. Kim, P., 2017. Matlab Deep Learning: With Machine Learning, Neural Networks and Artificial Intelligence. 1st Edn., Apress Publishing Company, New York, USA.,.
21. Oliveira, A.P., 2008. Prediction model of mechanical properties of hot-rolled structural beams: An approach in artificial neural networks. Master's Thesis, Federal University of Minas Gerais, Belo Horizonte, Brazil.
22. Bre, F., J.M. Gimenez and V.D. Fachinotti, 2018. Prediction of wind pressure coefficients on building surfaces using artificial neural networks. Energy Build., 158: 1429-1441.
23. Sari, Y.D. and M. Zarlis, 2018. Data-driven modelling for decision making under uncertainty. IOP Conf. Ser. Mater. Sci. Eng., Vol. 300, No. 1.
24. Ojha, V.K., A. Abraham and V. Snasel, 2017. Metaheuristic design of feedforward neural networks: A review of two decades of research. Eng. Appl. Artif. Intell., 60: 97-116.
25. Ge, Z., 2017. Review on data-driven modeling and monitoring for plant-wide industrial processes. Chemom. Intell. Lab. Syst., 171: 16-25.
26. Yin, S. and O. Kaynak, 2015. Big data for modern industry: Challenges and trends [point of view]. Proc. IEEE., 103: 143-146.
27. Sadati, N., R.B. Chinnam and M.Z. Nezhad, 2018. Observational data-driven modeling and optimization of manufacturing processes. Expert Syst. Appl., 93: 456-464.
28. Sorescu, A., 2017. Data-driven business model innovation. J. Prod. Innovation Manage., 34: 691-696.
29. Mosig, T., C. Lehmann and A.K. Neyer, 2020. Data-driven business model innovation: About barriers and new perspectives. Int. J. Innovation Technol. Manage., Vol. 1, 10.1142/S0219877020400179
30. McAfee, A., E. Brynjolfsson, T.H. Davenport, D.J. Patil and D. Barton, 2012. Big data: The management revolution. Harvard Bus. Rev., 90: 60-68.
31. Cosic, R., G. Shanks and S.B. Maynard, 2015. A business analytics capability framework. Australas. J. Information Syst.,, Vol. 19, 10.3127/ajis.v19i0.1150
32. Koohang, A. and M. Hatch, 2017. Leadership effectiveness in IT-centered organizations: Gender and levels of management. J. Comput. Inf. Syst., 57: 385-391.
33. Northouse, P.G., 2010. Leadership: Theory and Practice. 5th Edn., Sage, Thousand Oaks, California,.
34. Carillo, K.D.A., 2017. Let's stop trying to be sexy-preparing managers for the (big) data-driven business era. Bus. Process Manage. J., 23: 598-622.
35. Anderson, C., 2015. Creating a Data-Driven Organization: Practical Advice from the Trenches/Carl Anderson. 1st Edn., O'Reilly, Beijing, China,.
36. Anonymous, 2021. Difference between I-beam and H-beam steel. Vivadifferences.com, USA.