International Business Management
MEDWELL PUBLICATIONS

Artificial Intelligence: Risks and Opportunities

[1]Alisson Paulo De Oliveira and [2]Hugo Ferreira Tadeu Braga
[1]*Department of Metallurgical, Materials and Mining Engineering, Universidade Federal de Minas Gerais, Av. Antônio Carlos, 6627, Campus da UFMG, Pampulha, Escola de Engenharia, Brazil*
[2]*Innovation Center, Fundação Dom Cabral, Avenida Princesa Diana, 760, Alphaville, Lagoa dos Ingleses, Brazil*

Key words: Artificial neural nets, hot-rolled structural sections, prediction, autonomous direction, theory

Corresponding Author:
Alisson Paulo De Oliveira
Department of Metallurgical, Materials and Mining Engineering, Universidade Federal de Minas Gerais, Av. Antônio Carlos, 6627, Campus da UFMG, Pampulha, Escola de Engenharia, Brazil

Page No.: 236-243
Volume: 14, Issue 7, 2020
ISSN: 1993-5250
International Business Management

Abstract: This study aims to discuss the risks and opportunities involved in building predictive models based on artificial intelligence. Countermeasures are also proposed to minimize the risks involved in their adoption where reliability is a critical factor for user safety such as autonomous driving. For this, it is explored a real development of a predictive mathematical model, using industrial data in the steel industry. This development aimed to construct an empirical mathematical model to predict the mechanical properties (Yield Strength, YS) of hot rolled steel structural beams. Such model was based on rolling process variables and the chemical composition of steel. As a result of this research it was observed that the obtained data agreed with the expected metallurgical theory. The errors obtained between the estimated and the real values were greater for process conditions with lack of enough data. These results are associated with the risk of using artificial intelligence technology in critical applications and actions aiming at its improvement are proposed.

INTRODUCTION

Today there is a boom in the development of the most varied applications of artificial intelligence: In image recognition, autonomous driving capable cars and suggestion of medical treatments. Unfortunately, however, there are cases of failure leading to life-threatening accidents, according to Yadron and Tynan[1], bots assuming unwanted "personalities" as reported by Moreira, inappropriate medical treatment suggestions. according to Gnipper[2], among others.

According to Oliveira, etc., it is well established that the error in the training of artificial intelligence algorithms is smaller the greater the availability of data related to a given situation. So, necessarily, situations with little existing training data lead to poor learning with considerable response errors.

This study seeks to explore artificial neural networks, their definition and main characteristics, the learning mechanism, the risks arising from this process as well as opportunities for improvement aiming at better predictive results with fewer errors. It also seeks to suggest possible actions aimed at reducing this error and better response with response gains and reduced losses.

The artificial neural networks: The concept of artificial neural networks was inspired by the study of the human central nervous system. In the artificial neural network,

simple artificial nodes, known as neurons, processing elements or units, are connected to build a network that mimics the biological neural network of humans. The back propagation neural network is a well-known type of neural network with multilayer perceptual architecture with error back propagation for supervised learning and is particularly powerful for nonlinear prediction. It uses error back propagation to calculate the loss function gradient and modify model parameters (such as weights and activation limits) by the descending gradient method for each specific observation, according to Lee and Tsai[3].

Interactions between neurons characterize the transmission and processing of information and artificial neural networks have the advantage of strong adaptability, fault tolerance and noise prevention. These merits have made neural networks successfully applied in numerous fields, according to Ma *et al*.[4].

As an analytical tool artificial neural networks consist of numerous connections of artificial neurons or nodes. Each node represents a specific output function called the activation function. The connection between each two nodes represents a weighted value for the signal passing through the connection, called the weight which is equivalent to the memory (or learning) of the artificial neural network. Network output differs depending on network connection path, weighted value and the activation function. The network is usually an approximation of a certain algorithm or function or it can be an expression of a logical strategy. There are several types of artificial neural network algorithms including back propagation algorithms, radial base function algorithms and self-organizing ones. Back propagation neural networks can include three layers: the input layer, the hidden layer and the output layer as shown in Fig. 1. The model is trained by comparing the differences between network outputs and outputs reference and by minimizing its error by adjusting the connection weights. Then the reverse propagation neural network is constructed with acceptable errors by continuously adjusting its weights as reported by Ma *et al*.[5].

Some neurons can influence the environment by triggering actions. Learning is about finding weights that cause the neural network to exhibit desired behavior such as driving a car. Depending on the problem and how neurons are connected such behavior may require long causal chains of computational stages where each stage transforms (often in a nonlinear way) the aggregate activation of the network. So-called deep learning concerns the accurate assignment of weights through many similar stages as explored by Schmidhuber[6]

In contrast to procedural programming, the performance of an artificial neural network refers to its loss. This value, calculated across the entire training data

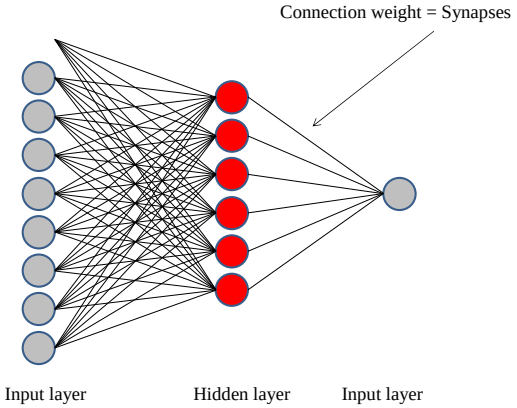

Fig. 1: Basic architecture of a three-layer reverse propagation neural network, designed by the researchers

Fig. 2: Information exchange in synapses through neurotransmitters as illustrated by Castrounis[7]

set for all inputs and possible outputs, reflects how well the weights that make up the trained network provide accurate predictions. Attempting to improve the accuracy of an artificial neural network is not about nullifying code errors but about finding ways to minimize total loss. This involves strategies such as the descending gradient that calculates the derivative of the loss function with respect to the parameters, then following that derivative, stepping backward and repeating until it reaches the base of the curve, indicating a minimum where the loss function is at its lower value. Or put another way, machine learning error is the crucial insight that allows us to modify the problem-solving process into one of numerical optimization as put by Collins[8].

Axons which branch into collaterals, receive signals from the cell body and lead them through the synapse to the dendrites of neighboring neurons as seen by Basheer and Hajmeer[9]. An artistic illustration of the signal transfer between two neurons across the synapse is shown in Fig. 2. An impulse, in the form of an electrical signal, travels inside the dendrite and through the cell body toward the synapse's presynaptic membrane.

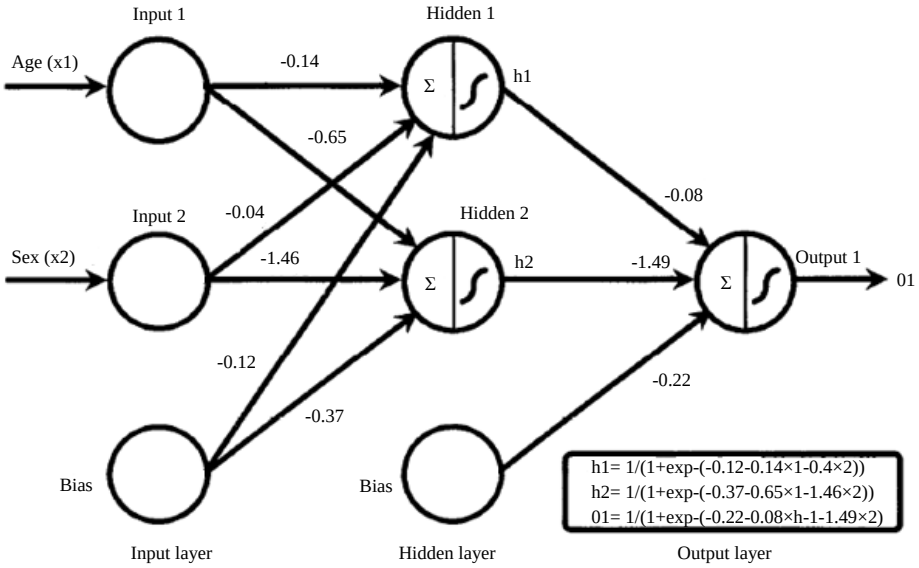

Fig. 3: Diagram of an artificial neural network trained to predict the likelihood of a patient dying from a hypothetical disease based on their age (x1) and gender (x2) as indicated by Tu[10]

Once the membrane is reached, a chemical neurotransmitter is released from the vesicles in amounts proportional to the strength of the newly arrived signal. The neurotransmitter diffuses into the synaptic breach toward the postsynaptic membrane and eventually into the dendrites of neighboring neurons, thereby, forcing them (Depending on the minimum value required for stimulus from the receiving neuron) to generate a new electrical signal from according to Basheer and Hajmeer[9].

The generated signal passes through the second neuron in the same manner as described above. The amount of signal that passes through a receptor neuron depends on the signal strength emanating from each of the feeder neurons (the ones that generate the signals), their synaptic strength and the minimum value required for stimulation of the receptor neuron. Because a neuron has many dendrites and synapses, it can receive and transfer many signals simultaneously. These signals may excite or inhibit neuron activity. This simplified signal transfer mechanism is the fundamental stage in the early development of neuro-computing and the operation of artificial neural networks. A fairly simplified analogy between an artificial and a biological neuron is based on the fact that the connections between the neurons represent the axons and dendrites, the weights of the connections represent the synapses (synaptic force) and the minimum value required for stimulation to represent activity in the body. according to Basheer and Hajmeer[9].

Figure 3 illustrates the operation of a three-layer artificial neural network constructed to predict the probability of an event (death from a hypothetical disease) as a function of two variables (age and gender of the patient) as indicated by Tu[10].

Neural networks can have multiple outputs. Nodes in the hidden layer (neurons) contain intermediate values which are calculated by the net. Each of the hidden and output nodes contains a function called the "Activation Function". Hidden nodes allow the network to model complex nonlinear relationships between input and output variables. Usually, each node in the input layer is connected to each node in the hidden layer and each node in the hidden layer is connected to each node in the output layer. In Fig. 1, there are two input nodes where the values of age, x1 and gender, x2 are entered into the network along with an adjustment weight (bias) which is the equivalent of the intercept term found in a model. Regression. At each node of the hidden layer, the input data multiplied by the respective connection weights are summed together with the adjustment weight. The result is used to calculate the node output through an activation function. The activation of each hidden node, h1 and h2 is then multiplied by a second set of connection weights (Example: 0.8, 1.49) and the result added to the adjustment weight (Example: -0.22). A transformation of the balanced inputs to the output nodes is applied to determine the total output of the network, according to Tu[10].

Deep learning and machine learning: In a simple definition machine learning or deep learning refers to the use of an artificial neural network with multiple layers of hidden nodes between output and input,

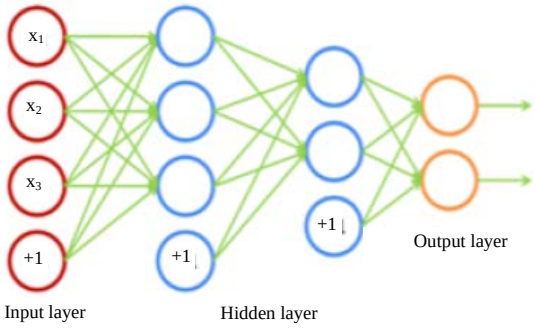

Fig. 4: Multilayer artificial neural network (Deep Network). Adapted from Nezhad *et al*.[11]

Fig. 4 where deep architectures are constructed by various levels of nonlinear operations, according to Nezhad *et al*.[11].

As shown in Fig. 4, the deep network has the same architecture as a traditional neural network but with a greater number of hidden layers. The main difference between a deep network and a traditional neural network is the algorithms developed for deep architecture training which are faster and yield stronger results. Deep learning includes representation learning algorithms that transform raw characteristics into high-level abstractions using a deep network composed of several hidden layers. In other words, deep learning applies computational approaches which have nonlinear multiple transformations, to train data representation through various levels of abstraction, according to Nezhad *et al*.[11].

Recently, the automatic extraction of features and representations has become an emerging area of human activity recognition research. To reduce dependency on human resources in engineering and the time spent on specific activities and applications, deep learning has become much sought after for this purpose. Deep learning uses appropriate machine learning techniques to present and model high-level representational features in sensor data using multiple layers of neural networks that represent low-to-high level hierarchical characteristics. Deep learning has become influential in research areas such as image and object recognition, natural language processing, machine translation, environmental monitoring and wearable and mobile sensors based on the recognition of human activity. Since, the improved implementation of deep learning in 2006, different methods have been developed and modified to solve a variety of challenging problems. These include boltzmann restricted machine, autoencoder, sparse encoding, convolutional neural networks and recurrent neural networks. Deep learning generally provides flexibility, robustness and performance enhancement with the power of multiple layers of neural networks, according to Nweke *et al*.[12].

Deep learning is essentially a repagination of artificial neural networks. In theory, a neural network with more than two layers, input and output can be classified as deep architecture. However, it is not just about the number of layers but about the idea of automated construction of more complex features in each step. This means that stacking other algorithms multiple times, using probabilities instead of class labels can be considered as deep learning as well. Back-propagation which has been around for decades, theoretically allows you to train a multi-layered neural network. Prior to technological advances in computational power, researchers did not have wide success in training more than two-layer neural networks simply because of the many calculations required to adjust network weights as seen by Khan and Yairi[13].

For conventional machine learning algorithms, it is difficult to extract well-represented features due to limitations such as dimensionality, computational bottlenecks and mastery requirements and specialized knowledge. Deep learning solves the problem of representation by constructing multiple simple features to represent a sophisticated concept. For example, a deep learning-based image classification system represents an object by describing its edges, tissues and structures in hidden layers. As increased training data availability, deep learning becomes more powerful and models based on this technology have solved many complicated problems with the help of hardware acceleration in computational time as explored by Yuan, etc.

Deep learning has made significant progress in a broad domain of machine learning: image classification, object recognition, speech recognition, language translation, voice synthesis. The online version of Go Master (AlphaGo) has beaten >50 elite players in the world. Recently, AlphaGo Zero surpassed its previous version without the use of human knowledge and a generic version, AlphaZero, reached a superhuman level within 24 h of cross-training in Chess, Shogi and Go. Driven by the emergence of Big-Data and hardware acceleration, deep learning requires less manual engineering resources and specialized knowledge. Data complexity can be extracted with a larger and more abstract level of representation from raw input characteristics as explored by Yuan, etc.

Deep learning is widely regarded as a “black box” technique where no information about the system is available beyond the data or if available, is not used in the model obtaining procedure. In this case, only system input and output data are used during identification. It is well known that it performs very well but with limited knowledge of the reasons why this happens. Many studies have been proposed to explain and interpret deep neural networks. By examining the opposing examples we can gain insights into the internal semantic levels of neural networks and find problematic decision boundaries which in turn help to increase the robustness and performance of

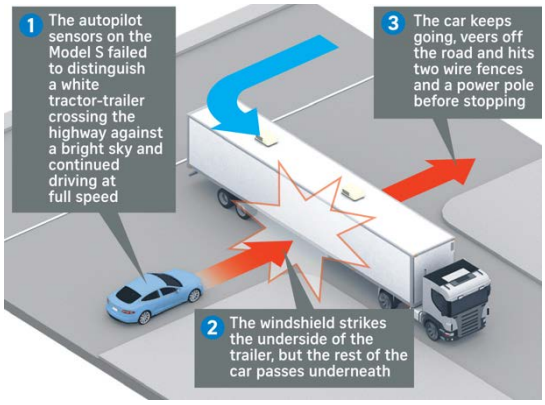

Fig. 5: Circumstances of an autonomous car accident, according to The Strait Times[14]

neural networks and thus, improve the interpretability of what occurs during the solution of neural networks. problems, according to Yuan, etc.

Typically, deep learning algorithms become better at certain tasks by filtering out huge amounts of data that humans have classified with the correct answers. This has allowed certain algorithms to reach accuracies well above 90% in certain training datasets that involve tasks such as correctly identifying human faces on Facebook or choosing the correct Google translation phrase from, say, Chinese and English. In such cases it is not the end of the world if a friend gets a misidentification or an esoteric phrase is translated incorrectly as explored by Yuan, etc.

But the consequences of errors increase dramatically as technology companies start using deep learning algorithms in applications such as two-ton machines moving on high-speed highways. A wrong decision by autonomous artificial intelligence can lead the car to collide with the guardrail, collide with another vehicle or run over pedestrians or cyclists. Government regulators will want to know for sure that autonomous cars can meet a certain level of safety and test data sets may not discover all those rare remote cases that could lead to an algorithm catastrophically failing as verified by Hsu[15].

For example, a Google autonomous car recently crashed into a bus because it hoped that it would prefer under a rare set of conditions but the bus did not. A Tesla autopilot vehicle collided with a truck because this system failed to recognize the truck as an obstacle due to its "white color against a bright sky" and high ride height. Such remote cases were not part of the set. Tesla and Google tests and thus never appeared during the test stages as Pei *et al.*[16]. Figure 5 illustrates this occurrence.

Debugging neural networks in autonomous vehicles involves rather tedious or random methods. A random test involves the creation by human researchers of test images that are fed into neural networks until they make a wrong decision. A second approach, called adversarial testing, can automatically create a sequence of test images by subtly adjusting an image until it causes the artificial neural network to trip, according to Hsu[15].

MATERIALS AND METHODS

In the development of a predictive model for industrial use, a structural steel beam with a nominal thickness of 11.0 mm was chosen as it has the following characteristics:

- Sampling for mechanical tests done on the flange which implies less variability of results
- Profile I with high number of tests performed in ASTM A572-50. For the present study, a total of 461 result sets were used

The following information was used in building the mechanical properties prediction model:

- Chemical composition of steel
- Final flange rolling temperature
- Measured thickness of the specimen from the profile flange (The thickness of the flange is the same as that of the specimen)
- Yield strength

After analyzing the available data, it was observed that 11 of the analyzed chemical elements had no result values indicated in all records, thus, being discarded. Because they are residual elements (Ni, Co, Ca, Ti, B, W, Zr, As, Sb, Te and Pb), it is considered that there is no impact on the prediction model.

Statistical analysis of database variables: The following analyzes and actions were performed:

- Correlation analysis between the various input variables and the outputs
- Statistical characterization of input data
- Data Processing Range: +/-3 standard deviations
- Elimination of outliers

These techniques were used to eliminate the presence of data that could compromise the reliability of the database.

Prediction model development: The type of artificial neural network most used in both general applications and mechanical properties prediction is reverse propagation.

Table 1: Summary of characteristics of artificial neural networks

Characteristics	Criteria	Command
Structural steel	Thickness, t, = 11.0 mm	-
Partition of data set	Training set = 75% Validation set = 25%	RANPERM
Normalization	-	-
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	BT	-
Number of hidden layers	1	-
Size of hidden layer	6	-
Net training mode	6	-
Normalization	10	-
Net weigh initialization	-	TRAINBR

An artificial neural network of the reverse propagation type consists of an input layer (input variables for the problem) an output layer (dependent variables) and one or more hidden layers (responsible for interpreting the data).

Table 1 shows the main parameters used to build the model. The development of the model followed the methodology cited by Basheer and Hajmeer[9] and recommendations of the MATLAB program as seen in The Mathworks[17].

The next step in the development of the artificial neural network was the simulation of mechanical property values from the input dataset. These results were compared with the real mechanical properties data and then calculated the percentage errors of each sample and the average percentage error of the simulation.

From the simulated values and measured real values, a graph was elaborated where it is possible to compare these values for the same sample which allows to graphically evaluate the error obtained as a function of this sample and its location in the distribution of the real data.

RESULTS AND DISCUSSION

Selection of variables of artificial neural networks: Correlation analysis was performed between the output variables and the input variables (chemical composition and rolling variables). The choice of the input variables was made based on the correlation analysis between these and the output variables, involving the elaboration of the correlation matrix, the evaluation of the correlation coefficients obtained and the elimination of the chemical elements considered residual, not intentionally added to the steelmaking process. The following variables were added to the model:

- Chemical composition: only the elements C, Mn, Si, S, Cr, Nb and N
- Process variables: final temperature and total rolling reduction

Table 2: Input and output variables included in the model

Variables	Description
Y1	YS (MPa)
X1	C (%)
X2	Mn (%)
X3	Si (%)
X4	S (%)
X5	Cr (%)
X6	Nb (%)
X7	N_2 (%)
X8	Final rolling temperature (°C)
X9	Total rolling reduction (%)

After this step and the discrepant points were eliminated, a database with 444 occurrences was obtained with the variables indicated in Table 2.

Training of artificial neural networks: Figure 6 shows the evolution of Sum of Squared Errors (SSE) for neural network training data for Yield Strength (YS) prediction. Figure 7 compares the measured and estimated resistance limit values for the 111 samples (25% of the database) used in the RNA validation tests and Fig. 8 shows histogram of the measured output, the Yield Strength.

As the histogram in Fig. 4 shows, the central region concentrates most occurrences of learning data. At the tails, due to the absence of more data, as in any normal distribution, learning does not occur as in the central region which is confirmed by the largest and smallest error of the estimated data (Simulated), respectively. For example as in Fig. 3: for samples 36 and 75 which have actual results located in the central region of the histogram as shown in Fig. 4, there is a small prediction error between the estimated and actual results. For samples 23 and 57 whose real results are the distribution extremes, normalized maximums and minimums, the estimated values show greater error in relation to the real values. And this is true for most 111 samples with results corresponding to the ends of the histogram. The real results farthest from the center with normalized amplitude around 0.50, always present estimated values with greater error, usually tending to the center of the

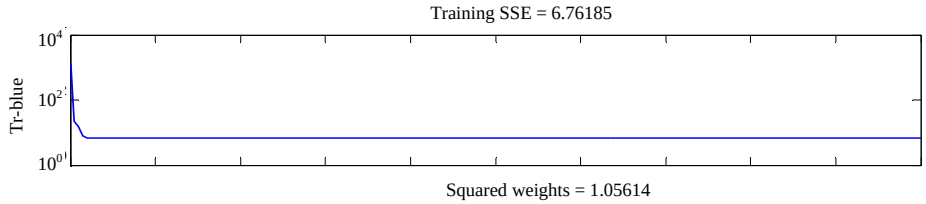

Fig. 6: Evolution of Sum of Squared Errors (SSE) for neural network training data for YS prediction

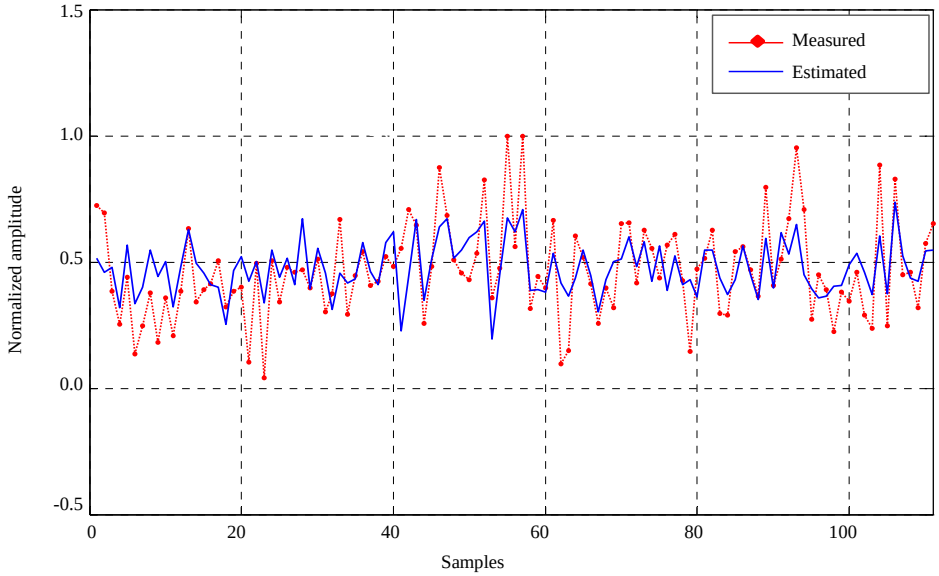

Fig. 7: Artificial neural networks validation results for YS prediction; Comparison between the measured and the estimated steel variable

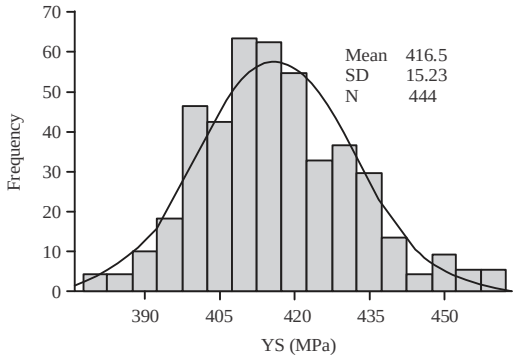

Fig. 8: Histogram of the actual values of the yield strength

distribution. In other words, the model based on artificial neural networks has reduced error in the central region of the data distribution and considerable error in the ends of this distribution. It is as if the model were not able to predict boundary situations or remote cases which leads to a catastrophic failure of the algorithm as discussed in the study. Other modeling work also showed similar behavior. As seen in above study where autonomous vehicle accidents under rare environmental conditions were discussed, we have that the example shown in the study is similar.

The tails of the histogram correspond to the rare conditions in which there was not enough neural network training to learn an important feature. Due to this fact, the error found in this region is significantly higher compared to the central region of the histogram, as can be seen in Fig. 4.

A possible solution to increase learning and reduce the prediction error is to use a step of design of experiments where data corresponding to the ends of the histogram are intentionally generated and thus provide greater learning. Thus, we would have less error of the model and in the case of autonomous car, less possibility of accidents.

For the practical example discussed in the study, just design the experiment design, so that, the maximum and minimum regions of the input variables (Table 2) have enough occurrences of

training data. For this, the process must be modified in its origin, thus, forcing the occurrence of this data.

CONCLUSION

This study discussed artificial neural networks as well as the back propagation training method, showing in detail what their working principle is like. Deep learning technology as well as the risks of its use in a real case involving autonomous driving were also seen. The construction of a predictive model using industrial data for physical characteristics of steel products as well as the errors obtained in this process and their association with the amount of data available for learning was explored. A methodological approach to minimize errors associated with regions with little real data availability has been suggested to improve the learning process. This study is expected to be useful for your readers to learn.

REFERENCES

01. Yadron, D. and D. Tynan, 2016. Tesla driver dies in first fatal crash while using autopilot mode. The Guardian, Kings Place, London. https://www.theguardian.com/technology/2016/jun/30/tesla-autopilot-death-self-driving-car-elon-musk
02. Gnipper, P., 2018. Watson, IA da IBM, fez recomendacoes inseguras para tratar cancer. HostGator, Houston, Texas, USA. https://canaltech.com.br/inteligencia-artificial/watson-ia-da-ibm-fez-recomendacoes-inseguras-para-tratar-cancer-118930/
03. Lee, C.Y. and T.L. Tsai, 2019. Data science framework for variable selection, metrology prediction and process control in TFT-LCD manufacturing. Rob. Comput. Integr. Manuf., 55: 76-87.
04. Ma, J., Y. Ma and C. Li, 2019a. Infrared and visible image fusion methods and applications: A survey. Inf. Fusion, 45: 153-178.
05. Ma, J., D.W. Sun, H. Pu, Q. Wei and X. Wang, 2019b. Protein content evaluation of processed pork meats based on a novel single shot (snapshot) hyperspectral imaging sensor. J. Food Eng., 240: 207-213.
06. Schmidhuber, J., 2015. Deep learning in neural networks: An overview. Neural Networks, 61: 85-117.
07. Castrounis, A., 2016. Artificial intelligence, deep learning and neural networks, explained. KDnuggets, Chestnut Hill, Massachusetts, USA. https://www.kdnuggets.com/2016/10/artificial-intelligence-deep-learning-neural-networks-explained.html
08. Collins, J., 2017. Glossary of deep learning: Error. A Medium Corporation, Spanish, French. https://medium.com/deeper-learning/glossary-of-deep-learning-error-1f70d9bb88e9
09. Basheer, I.A. and M. Hajmeer, 2000. Artificial neural networks: Fundamentals, computing, design and application. J. Microbiol. Meth., 43: 3-31.
10. Tu, J.V., 1996. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. J. Clin. Epidemiol., 49: 1225-1231.
11. Nezhad, M.Z., N. Sadati, K. Yang and D. Zhu, 2019. A deep active survival analysis approach for precision treatment recommendations: Application of prostate cancer. Expert Syst. Appl., 115: 16-26.
12. Nweke, H.F., Y.W. Teh, G. Mujtaba and M.A. Al-Garadi, 2019. Data fusion and multiple classifier systems for human activity detection and health monitoring: Review and open research directions. Inf. Fusion, 46: 147-170.
13. Khan, S. and T. Yairi, 2018. A review on the application of deep learning in system health management. Mech. Syst. Signal Proc., 107: 241-265.
14. The Strait Times, 2016. US probes second suspected Tesla Autopilot crash. The Strait Times, Singapore. https://www.straitstimes.com/world/united-states/us-probes-second-suspected-tesla-autopilot-crash
15. Hsu, J., 2017. A new way to find bugs in self-driving AI could save lives. IEEE Spectrum, New York, USA. https://spectrum.ieee.org/tech-talk/artificial-intelligence/machine-learning/better-bug-hunts-in-selfdriving-car-ai-could-save-lives
16. Pei, K., Y. Cao, J. Yang and S. Jana, 2017. Deepxplore: Automated whitebox testing of deep learning systems. Proceedings of the 26th International Symposium on Operating Systems Principles, October 2017, ACM, Shanghai, China, pp: 1-18.
17. The Mathworks Inc., 2008. Documentation about neural network toolbox. The Mathworks Inc., Natick, Massachusetts, USA.