BAYESIAN REGULARIZERS OF ARTIFICIAL NEURAL NETWORKS APPLIED TO THE RELIABILITY FORECAST OF INTERNAL COMBUSTION MACHINES IN THE SHORT-TERM.

Ítalo Rodrigo Soares Silva
Student of the Graduate Program in Engineering, Process, Systems and Environmental Management (PPG-EGPSA), ITEGAM, Brazil.
Email: italo.computation@gmail.com
ORCID: https://orcid.org/0000-0002-2827-1078

Manoel Henrique Reis Nascimento
PhD in Electrical Engineering from the Graduate Program in Engineering, Process, Systems and Environmental Management (PPG-EGPSA), ITEGAM, Brazil.
Email: hreys@bol.com.br
ORCID: https://orcid.org/0000-0003-4688-6751

Milton Fonseca Júnior
PhD in Electrical Engineering from the Graduate Program in Engineering, Process, Systems and Environmental Management (PPG-EGPSA), ITEGAM, Brazil.
Email: milton.fonseca.jr@gmail.com
ORCID: https://orcid.org/0000-0001-6001-3098

Ricardo Silva Parente
Student of the Graduate Program in Engineering, Process, Systems and Environmental Management (PPG-EGPSA), ITEGAM, Brazil.
Email: ricardosilvaparente@gmail.com
ORCID: https://orcid.org/0000-0003-3096-9069

Paulo Oliveira Siqueira Júnior
Student of the Graduate Program in Engineering, Process, Systems and Environmental Management (PPG-EGPSA), ITEGAM, Brazil.
Email: paulojunior051996@gmail.com
ORCID: https://orcid.org/0000-0001-8071-3972

Jandecy Cabral Leite
PhD in Electrical Engineering from the Graduate Program in Engineering, Process, Systems and Environmental Management (PPG-EGPSA), ITEGAM, Brazil.
Email: jandecycabral@hotmail.com
ORCID: https://orcid.org/0000-0002-1337-3549

Abstract

Predictive as well as preventive maintenance are tools of maintenance programs that aim to increase or maintain the life expectancy of an equipment through computational techniques and tools. Bearing in mind that the power generation industry has a high maintenance rate with machines and / or electric generators stopped, this research aims to develop a computational model for predicting the Reliability Key Performance Indicator (KPI) to identify how available the equipment will be in a time span of 22 days, for this the methodology to be used will be based on analyzes and tests of artificial neural network (ANN) architectures using the Bayesian Regularizers training algorithm, alternating the transfer functions in the layers hidden to find the best state of convergence and the minimum Root Mean Square Error (RMSE) value calculated between the real and simulated outputs. According to the results obtained by the training, validation and test steps, the algorithm presented a RMSE rate of 0.0000104202 and a 99.9% correlation between the real and simulated values, thus the model is able to identify which machine will have the greatest efficiency and less efficiency within the defined time span.

Keywords: Reliability, RNA, Bayesian Regularizers, UTE;

1. Introduction

One of the elements that cause financial impacts in the sectors of commerce and industry is the maintenance of electric machines and / or generators (CORRÊA, 2020; SALLES *et. al.,* 2020; RUIZ-HERNÁNDEZ *et. al.,* 2020) that require large workforce for carrying out energy supply procedures and not having work breaks. Among the concerns of this sector, the supply of energy stands out (SÁNCHEZ *et. al.,* 2020), as occurs in companies of essential services such as hospitals, clinics and supermarkets causing eventual irreversible damage or even loss of life in the case of hospitals when there is a lack of energy supply (CHINI *et. al.,* 2020).

For this, the other types of maintenance become trivial to guarantee the functioning of the equipment, among which qualified professionals who adapt or adopt methodologies of different procedures to perform in the necessary environment (AGNESE, 2020) perform the methods of prevention, prediction and correction. One of the performance indicators used in maintenance management according to NBR-5462 is Reliability, which refers to the probability of proper functioning within a certain period of the elements involved in the production chain.

One of the methods used to identify faults in electric generators that includes internal combustion machines is the vibration analysis (ROCHA, *et. al.,* 2020; ZINGONI, 2020), the method consists of monitoring the target machine by defining the measurement ranges and parameters used for data collection. However, with the advancement of science and technology, computational tools gain space in these scenarios in order to optimize and improve the quality with which the processes are performed, but what stands out most in these procedures is the appropriate use of computational intelligence techniques (SOUZA, 2020; ABDULRAHMAN *et. al.,* 2020; BAROROH *et. al.,* 2020) to guarantee, through mathematical models disseminated in academia and science, the results presented by the tool.

With this, studies with Machine Learning, Deep Learning, Pattern Recognition, Data Processing, Pattern

Classification, Optimization Techniques (RIGHETTO, 2020; ABDULRAHMAN *et. al.,* 2020; BAROROH *et. al.,* 2020) and many others are growing as the demands for complexity in these scenarios increase. An example that contemplates the scenario of maintenance of machines with the use of AI (Artificial Intelligence) is the detection of failures (CARDOSO, 2020; ARUNTHAVANATHAN *et. al.,* 2020).

Therefore, the use of these elements becomes more evident and tends to raise the level of complexity according to the need that companies find to adapt these methods or tools to the scenario used and obtain satisfactory results. With this, the present research demonstrates a focus on the scenario of maintenance of machines when using the group of electric generators of Thermoelectric Power Plants as study object, the study and implementation of the Artificial Neural Network Technique for Prediction of the Group's Reliability Index of engines.

2 Literature Review

2.1 Predictive maintenance

Predictive maintenance (HAN *et. al.,* 2021; SILVA *et. al.,* 2019) becomes a necessary method since it is possible to prevent the stoppage of machines in a production process through indicators offered by monitoring systems, identify the small irregularities that can evolve to large failures early and thus allow for correction (AYVAZ and ALPAY, 2021; TIAN, LIU and SHU, 2021). Some methods for performing equipment-monitoring diagnostics are used in the literature (SCHWENDEMANN, AMJAD AND SIKORA, 2021; MANHERTZ and BERECZKY, 2021; LUGHOFER and SAYED-MOUCHAWEH, 2019), such as:

Vibration Analysis: Through analysis of machines excited by dynamic efforts, vibration sensors at defined points in the machine, the vibration registers are captured (FONSECA-JUNIOR *et. al.,* 2015). According to the extent to which the components of an equipment start to fail, the frequency and amplitude of vibration begin to change and with the analysis of the spectrum applied to the system it is possible to identify whether any component has its integrity compromised (MORO, 2020; YU, FENG and LIANG, 2021).

Thermography: This method is based on the detection of infrared radiation emitted naturally by bodies with temperature proportional to the intensity emitted from the equipment, with this, it is possible to identify regions or points where the temperature is altered and obtain information about the state of the machine (MEIßNER *et. al.,* 2021; LUGHOFER and SAYED-MOUCHAWEH, 2019; FONSECA-JUNIOR *et. al.,* 2015).

Cracks analysis: Uses magnetic particle test methods, deviate from their trajectory when finding a superficial or subsurface discontinued, this allows to identify points of non-conformity and apply the necessary repairs (CHEN *et. al.,* 2021; LUGHOFER and SAYED-MOUCHAWEH, 2019).

Thickness measurement: This method uses ultrasound as a non-destructive test, commonly used in industries to detect discontinuities in the entire volume of the material, the process consists of making the ultrasonic wave emitted by a transducer travel through the analyzed material, with this being verified the echoes received back, so it is possible to identify internal flaws or thicknesses (EL-ADAWY *et. al.,* 2021; LUGHOFER and SAYED-MOUCHAWEH, 2019).

2.2 Reliability in maintenance

According to NBR 5462, reliability is the possibility for an item, equipment, machine or system to perform its function within a certain period of time within a project, what is expected with the method is to analyze the degree of confidence of the equipment and with this determines metrics of use or maintenance techniques that act in the continuous improvement of the equipment (LU *et. al.,* 2021).

This measurement is commonly made based on a history of equipment performance and its estimate of future operation, which can be measured: Increased machine life, reduced maintenance costs, improved operational performance, agility and consistency of the teams techniques (ZOU *et. al.,* 2021; LUGHOFER and SAYED-MOUCHAWEH, 2019). The reliability parameter to be calculated from this historical observation can be expressed by equation 1:

$$\lambda_i^x = \frac{N_{fi}}{t_i^x . n} \tag{1}$$

Where N_{fi} is the number of failures before the i-th point of failure, $t_i^x$is the failure time of the ith point of failure for subsystem x, and n is the number of failures of the subsystem (BAI *et. al.,* 2020). The failure rate for each failure point of subsystem x can be expressed by equation 2:

$$\lambda^x = [\lambda_1^x, \lambda_2^x, \lambda_3^x \dots \lambda_i^x, \dots \lambda_{n-1}^x, \lambda_n^x]^t \tag{2}$$

For Bai et. al. (2020) in relation to fault investigation techniques, one of the most used is the bathtub curve, in which it is possible to analyze the equipment's useful life, a series of combinatorial radial basis functions for RBF are used to approximate functions complex or difficult calculations, expressed by equation 3.

$$\hat{y}(x) = \sum_{i=1}^{n} \beta_i f(\|x - x_i\|) = f(x)^T \beta \tag{3}$$

Where $\hat{y}(x)$ is the prediction response vector, x, β is the radial base coefficient vector β_i is the i-th component $\beta, f(x)$ is the RBF vector $f(\|x - x_i\|)$ is the i-th component of $f(x), r = \|x - x_i\|$ is the Euclidean distance between two vectors (LUGHOFER and SAYED-MOUCHAWEH, 2019; BAI *et. al.,* 2020). The same equation can be rewritten for the failure rate equation being expressed by equation 4:

$$\hat{\lambda}^x(t) = \sum_{i=1}^{n} \beta_i f(\|x - x_i\|) = f(x)^T \beta \tag{4}$$

Where $\hat{\lambda}^x(t)$ is the failure rate of subsystem x at time t. The unit's reliability is defined as its cumulative probability of success, thus the reliability function $R(t)$ is given by equation 5:

$$R(t) = \frac{n_s(t)}{n_s(t) + n_f(t)} = \frac{n_s(t)}{n_0} \tag{5}$$

The distribution function $F(t)$is the complement of $R(t)$ expressed by equation 6:

$$R(t) = 1 - F(t) = 1 - \int_0^t f(u)du = \int_t^{+\infty} f(u)du \tag{6}$$

Thus, the reliability function indicates the probability that an item / equipment will be successful in its operation, characterized by the absence of failures in a period of time (LUGHOFER and SAYED-MOUCHAWEH, 2019; BAI *et. al.,* 2020). In order to measure reliability by means of probabilistic calculations, it is necessary to design a future scenario. However, it is necessary to have a survey of the average time between equipment failures, equipment failure rate and to know in advance what will be projected for the reliability calculation, in this case. if the number of projection days (SOLTANALI *et. al.,* 2021).

2.3 Bayesian regularizers in machine learning

The Bayesian inference can also be used to select the best structures or hypotheses $H = \{H_1, H_2, \dots, H_k\}$. According to Barbosa and Ferreira (2020) with the Bayes rule, the a posteriori probability distribution $p = (H_h|Y)$ of the H_h hypothesis is given by, expressed by equation 7:

$$p = (H_h|Y) = \frac{p = (H_h|Y)\ \mathrm{p}(H_h)}{p(Y)} \tag{7}$$

Where:

$p(Y)$ = is a normalization factor

H_h = a priori equiprobable;

$p(H_h|Y)$ = model evaluator

The a priori probability $p(Y)$ starts from the premise that weights should initially assume values close to zero in order to avoid training saturation in analogy to the normalization applied to the input and output signals (FERREIRA and BARBOSA, 2020; FERREIRA, DE SOUZA and DO COUTTO FILHO, 2020). To define a probability, let δ_k be the output of the k-th output neuron of the Multilayer Perceptron (MLP), the equation x represents the weights that connect the hidden layer neurons to the k-th output neuron, expressed by equation 8.

$$\delta_k = \emptyset_{saida}\left\{\sum_{j=1}^{m} \varpi_{kj}\phi_{oculta}\left(\sum_{i=1}^{n} w_{ji}x_i + b_j\right) + b_k\right\} \tag{8}$$

The bias of this neuron, $b_k \in \mathbb{R}$ o bias of the k-th output neuron, $\emptyset_{oculta}(.)\colon \mathbb{R} \longrightarrow \mathbb{R}$ the sigmoidal function of activation of the neurons of the hidden layer and $\emptyset_{saida}(.)\colon \mathbb{R} \longrightarrow \mathbb{R}$ the function of activation of neurons in the output layer. Thus, the vector $\underline{w} \in \mathbb{R}^m$, $\underline{w} = [\underline{w}_s'\ \underline{w}_1' \dots \underline{w}_j'\ b\ b_1 \dots b_j]^t$ training without the need for a validation set (FERREIRA and BARBOSA, 2020; FERREIRA, DE SOUZA and DO COUTTO FILHO, 2020).

The principle of maximizing the evidence applied to the parameters $\underline{w}$ giving rise to the functional $S(\underline{w})$ to be minimized for estimation of $\underline{w}$ applied the hypotheses $H = \{H_1, H_2, \dots, H_k\}$ to calculate ln $p = (H_h|Y)$ is also applied to the hyperparameters $\alpha_i\ e\ \beta$ giving rise to an iterative algorithm (LEOCÁDIO and FERREIRA, 2012). For Leocádio and Ferreira (2012) the Bayesian inference based on maximizing the evidence applied to the development of MLPs can be summarized by the algorithm:

Step 1: specify the minimum number (N_{min}) and the maximum number (N_{max}) of neurons in the hidden layer and make the number of neurons $m = N_{min}$;

Step 2: manipulate proof variables in a n-dimensional vector of inputs. In cases of binary inputs (n + 2), if they are continuous (n + 1);

Step 3: do l =0 and initialize $\underline{w}(l) = [\underline{w}_1(l), \dots, \underline{w}_{n+3}(l)]^t, \underline{\alpha}(l) = [\underline{\alpha}_1(l), \dots, \underline{\alpha}_{n+3}(l)]^t\ e\ \beta(l)$.

Step 4: Using error back propagation, minimize $s(\underline{w}) = \frac{\beta}{2}\sum_{j=1}^{n}[d_j - f(\underline{x}_j, \underline{w})]^2 + \frac{1}{2}\sum_{i=1}^{n+3}\alpha_i \sum_{l=1}^{M} w_{il}^2$ about $\underline{w}(l)$ to obtain $\underline{w}(l+1)$.

Step 5: Calculate $\alpha_i(l+1), \beta(l+1)\ e\ \gamma_i(l+1)$;

Step 6: Do l = l + 1 and return to step 4 until convergence. After convergence, proceed to the next step.

Step 7: Separate the hyperparameters $\alpha_1's$ related to continuous inputs and the hyperparameters $\alpha_j's$ into binary inputs into two subsets.

Step 8: For each list, select the entries with $\alpha_i < \alpha_{ref}$ represents the hyperparameter associated with the proof variable.

Step 9: Repeat steps 4 to 6 using only the entries selected in steps 8, with n answering for the number of entries selected to obtain the trained model H_m

Step 10: Calculate the logarithm of the evidence for the hypotheses (number of neurons in the hidden layer) H_m.

Step 11: If $m = N_{max}$ go to step 12, otherwise do m = m + 1 and return to step 2.

Step 12: select the hypothesis H_k with more evidence ln $p(\underline{Y}|H_h)$ to make the predictions.

Each RNA model has characteristics and purposes, it is worth noting that the use of these models depends on the architecture and the learning process that must be balanced with a training algorithm. Learning occurs when the neural network reaches a generalized solution to a class of problems (CABEZA *et. al.,* 2018). Among the ways to learn a neural network there are those that consist of: error correction, competition, Hebrew models and learning machines (ARABI BULAGHI *et. al.*, 2020).

However, this alone is not enough to have the best neural network model, this is due to the various nonlinear applications, the activation functions do this intuitively by creating learning models that relate dependent and independent variables. Some examples of the activation function are shown in table 1.

Table 1: Activation functions.

Initials	Function	Expression
Sigmoide	Sigmoide	$\sigma = \frac{1}{1+e^x}$
TanH	Hyperbolic Tangent	$tanh(x) = 2\sigma(2x) - 1$ $tanh'(x) = 1 - tanh^2(x)$
ReLU	Rectified Linear Unit	$ReLU(x) = \max\{0, x\}, sendo \begin{cases} 1, se\ x \geq 0 \\ 0, c.c \end{cases}$
ELU	Exponential Linear Unit	$ELU(x,\alpha) = \begin{cases} x, se\ x \geq 0 \\ \alpha(e^x - 1), se\ c.c \end{cases}$ $ELU'(x,\alpha) = \begin{cases} 1, se\ x \geq 0 \\ ELU(x,\alpha) + \alpha, se\ c.c \end{cases}$
Leaky ReLU	Leaked Rectified Linear Unit	$LeakyReLU(x,\alpha) = \max\{\alpha x, x\}$ $LeakyReLU'(x,\alpha) = \begin{cases} 1, se\ x \geq 0 \\ \alpha, se\ c.c \end{cases}$

Depending on the number of iterations that have been defined for a neural network, the combination of these methods can reduce the prediction time or increase if not used correctly.

3. Materials and methods

In order to carry out this research, it was necessary to use a database containing the main failures of a group of internal combustion engines provided with vibration analysis in Thermoelectric Plants. The analysis period consists of January to December 2019, the data set includes date, time, downtime, system that failed, subsystem that was affected and description of the failure. Figure 1 illustrates a flow with the stages of development.

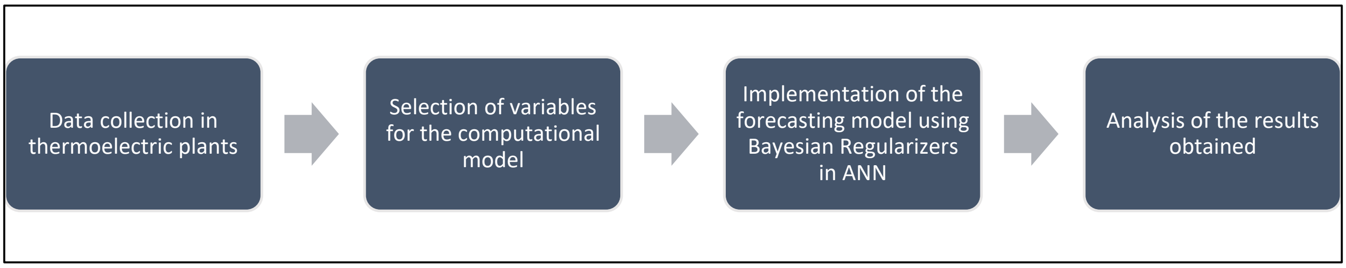

Figure 1: Development stages.

3.1 Step 1: Data collection in thermoelectric plants

For data collection, a survey of failures in internal combustion machines in thermoelectric plants was carried out in a period from January to December 2019. Table 2 shows the days analyzed during the period.

Table 2: Distribution of analyzes by month.

Month	Days analyzed	Operating hours
Janeiro	9	216
Fevereiro	26	624
Março	29	696
Abril	29	696
Maio	29	696
Junho	27	648
Julho	29	696
Agosto	27	648
Setembro	28	672
Outubro	21	504
Novembro	26	624
Dezembro	16	384
Total	**296**	**7104**

According to the analysis of the monthly samples they vary in an average of 24 days, where the hours of operation of the group of engines is given by equation 9:

$$HO = 24 * DA \tag{9}$$

Where:

HO = Hours of operation;

DA = Days analyzed;

24 = hours of the day;

The machines that were analyzed in that time are 4-stroke internal combustion engines, from Wartisila NSD

Corporation, model 18V46, with a nominal power of 15.75 MW, efficiency of 42.3%, length of 13.58m, width 5,347m, height of 5,488m and weight of 237 tons.

3.2 Step 2: Selection of variables for the computational model

The set of data provided by the vibration analysis allowed a table to be produced with the main variables that will be used in the forecasting model, where each one has importance and dependence for the calculation of Reliability, being:

1. **Total downtime**: total downtime of the machines per month;
2. **Total frequency**: number of occurrences of the machines per month;
3. **Days analyzed**: Number of days analyzed to provide failure data;
4. **Hours of operation**: total time of operation without considering failures;
5. **MTBF**: average time between failures;
6. **MTTR**: average repair time;
7. **Failure rate**: instantaneous failure rate within a period of time.

Finally, the **Reliability** output variable defined by a forecast time function to estimate this key performance indicator on a percentage scale.

3.3 Step 3: Implementation of the forecasting model using Bayesian Regularizers in ANN

For simulations of the computational model using the Bayesian Regularizers training algorithm in the learning process of the Neural Network, 5 input and 1 output variables were used. The network learning process is divided into 3 stages: training, validation and testing, where the data are separated by 70%, 15% and 15% respectively.

The machine that will generate the results provided by the simulations has the following configurations: 16GB of RAM, core i5 generation 10 processor with 2.50 GHz, 64-bit platform and 500GB SSD. Table 2 presents the configuration data of the neural network that were used to simulate the reliability of the motor group of the thermoelectric plants considering the Bayesian regularization.

Table 3: Configuration of the Bayesian model of RNA.

ALG	FC1	FC2	NC1	NC2	Epoch
Bayesian Regularization	Sigmoide	Sigmoide	24	50	24

Where:

ALG = Algorithm used;

FC1 = Layer 1 transfer function;

FC2 = Layer 2 transfer function;

NC1 = Number of neurons in layer 1;

NC2 = Number of neurons in layer 2;

Times = Number of iterations for the convergence of the algorithm;

MSE = Mean Square Error;

RMSE = Square Root of the Average Error

3.4 Step 4: Analysis of the results

The step for choosing the best configuration for the Bayesian regularization algorithm is an analysis of the lowest mean square error, where each network model is trained and a filter is made to identify the one with the lowest EMQ index. For each iteration performed, a network performance validation is performed, where a function is called that is responsible for simulating the network test result with the output vector, according to equation 10:

$$EQM = \sqrt{\frac{\sum_{i=1}^{n}(ra_i - rs_i)^2}{n}} \quad (10)$$

Where:
NDE = Mean Square Error;
n = number of elements of the output vector;
ra = target result;
rs = simulated result
Figure 2 illustrates the steps to obtain the best RNA model based on the minimum mean square error.

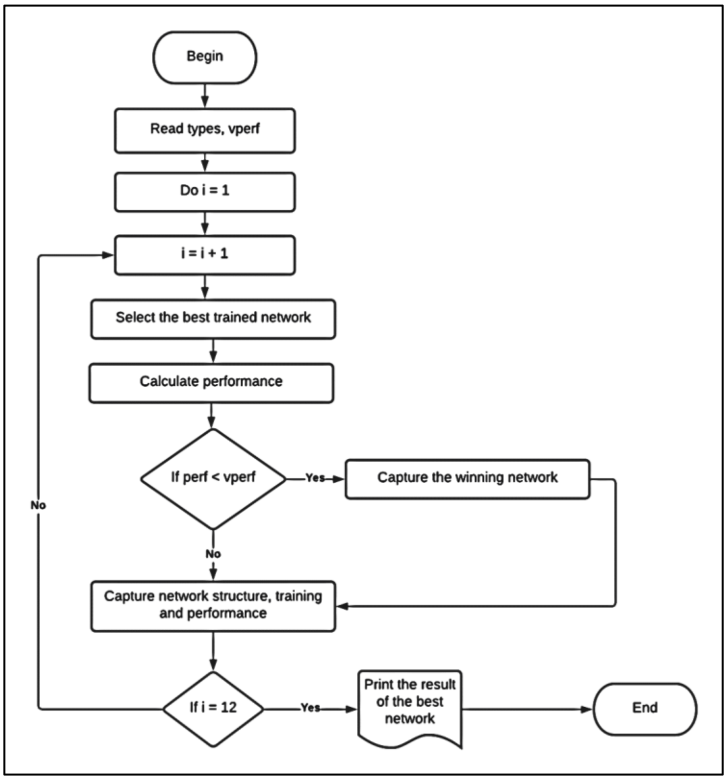

Figure 2: Best RNA selection algorithm.

For the evaluation of the winning Bayesian regularizer model, the configurations are alternated according to the need for neurons and the transfer function to escape the gradient explosion, thus obtaining better results of approximation of function.

4. Results and discussions

To find the data set necessary to feed the RNA input nxp matrix, some elements were pre-processed and the model was adjusted to a point where it was measurable, for example, in the case of considering an estimated assessment of the reliability of a group of 20 engines within a Thermoelectric Plant it was necessary to group the Total Downtime and the Total Frequency Time.

Table 4 presents the pre-processed data set for RNA in a period from January to December 2019, grouping the 20 TTP engines. These data are provided from the vibration analysis where the stop time calculation is performed for each equipment that failed, at the end of each month an accounting is made and grouped.

Table 4: TTP nxp array.

MOTOR	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ	TOTAL
1	0,00	3,0	2,37	0,00	1,92	0,00	0,00	7,30	1,8	2,68	0,0	1,3	**20,32**
2	0,00	0,0	0,45	8,98	0,85	3,87	6,33	1,30	0,0	0,70	9,3	0,0	**31,78**
3	0,00	5,9	9,52	0,00	5,20	0,00	0,45	0,70	0,0	0,00	0,0	0,9	**22,60**
4	4,13	7,4	2,60	20,53	2,85	3,43	1,17	4,35	2,0	0,00	1,8	0,0	**50,18**
5	0,00	0,0	1,98	0,00	0,00	0,00	0,65	0,00	0,0	0,00	0,0	2,1	**4,75**
6	0,00	7,8	0,00	1,92	2,52	4,97	0,00	0,00	4,5	5,10	3,2	1,0	**30,98**
7	0,00	0,0	0,00	0,45	3,43	3,00	3,20	0,00	5,0	0,00	0,5	1,9	**17,35**
8	8,92	0,0	0,00	0,00	0,00	0,70	0,00	6,85	0,0	4,52	0,0	0,0	**20,98**
9	0,00	2,2	1,92	1,08	1,92	0,00	0,45	0,00	1,9	2,13	2,7	0,0	**14,18**
10	0,00	0,0	2,00	3,30	1,80	3,93	0,00	1,45	0,0	1,60	1,4	2,2	**17,75**
11	7,37	1,5	2,00	0,00	0,00	0,00	6,25	7,92	1,3	0,00	7,8	0,9	**35,05**
12	3,43	3,9	0,00	0,85	9,35	7,85	0,00	1,98	0,0	0,00	0,0	0,0	**27,32**
13	0,00	2,0	12,38	3,25	0,00	2,93	6,85	0,00	1,3	0,00	2,0	7,4	**38,10**
14	14,15	0,8	0,00	0,00	3,92	6,10	0,00	1,75	5,8	3,67	0,9	3,3	**40,27**
15	0,00	0,0	0,00	0,00	2,23	2,00	7,00	12,03	0,0	3,20	0,0	1,5	**28,00**
16	3,85	1,5	0,85	8,95	1,33	0,00	0,70	1,75	5,2	1,53	0,7	0,0	**26,33**
17	0,00	3,5	5,00	1,92	5,62	4,10	2,00	0,63	0,0	1,75	6,2	0,0	**30,70**
18	0,00	1,9	0,00	5,85	0,45	0,00	0,85	3,95	0,0	1,98	0,0	0,0	**14,93**
19	0,00	1,0	12,82	2,73	4,10	0,00	3,20	3,08	2,4	0,00	0,0	0,0	**29,32**
20	2,93	0,1	1,72	11,08	1,75	5,17	0,45	1,92	2,6	0,00	1,9	6,7	**36,30**

Table 5 shows the pre-processed set of TF for RNA for a period from January to December 2019, grouping the 20 engines. The set represents the amount of frequency that the equipment failed during the month being grouped by engine, while table 6 shows the total downtime on a decimal hour scale.

Table 5: TF nxp array.

MOTOR	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ	TOTAL
1	0	1,0	2,00	0,00	1,00	0,00	0,00	2,00	1,0	2,00	0,0	1,0	**10**
2	0	0,0	1,00	3,00	1,00	3,00	4,00	1,00	0,0	1,00	3,0	0,0	**17**
3	0	2,0	2,00	0,00	2,00	0,00	1,00	1,00	0,0	0,00	0,0	1,0	**9**
4	1	2,0	2,00	3,00	2,00	2,00	1,00	3,00	1,0	0,00	1,0	0,0	**18**
5	0	0,0	1,00	0,00	0,00	0,00	1,00	0,00	0,0	0,00	0,0	2,0	**4**
6	0	3,0	0,00	1,00	2,00	4,00	0,00	0,00	3,0	2,00	1,0	1,0	**17**
7	0	0,0	0,00	1,00	2,00	3,00	0,00	0,00	2,0	0,00	1,0	1,0	**10**
8	1	0,0	0,00	0,00	0,00	1,00	0,00	1,00	0,0	1,00	0,0	0,0	**4**
9	0	2,0	1,00	2,00	1,00	0,00	1,00	0,00	1,0	1,00	2,0	0,0	**11**
10	0	0,0	2,00	2,00	3,00	3,00	0,00	1,00	0,0	2,00	1,0	1,0	**15**
11	2	1,0	1,00	0,00	0,00	0,00	3,00	2,00	1,0	0,00	4,0	1,0	**15**
12	3	1,0	0,00	1,00	3,00	4,00	0,00	1,00	0,0	0,00	0,0	0,0	**13**
13	0	2,0	2,00	1,00	0,00	2,00	1,00	0,00	1,0	0,00	1,0	5,0	**15**
14	1	1,0	0,00	0,00	2,00	2,00	0,00	1,00	2,0	3,00	1,0	1,0	**14**
15	0	0,0	0,00	0,00	1,00	1,00	3,00	4,00	0,0	1,00	0,0	1,0	**11**
16	1	1,0	1,00	2,00	2,00	0,00	1,00	1,00	2,0	1,00	1,0	0,0	**13**
17	0	2,0	2,00	1,00	3,00	1,00	1,00	1,00	0,0	2,00	4,0	0,0	**17**
18	0	1,0	0,00	2,00	1,00	0,00	1,00	2,00	0,0	1,00	0,0	0,0	**8**
19	0	1,0	2,00	2,00	1,00	0,00	1,00	2,00	1,0	0,00	0,0	0,0	**10**
20	1	1,0	2,00	3,00	1,00	2,00	1,00	1,00	2,0	0,00	1,0	2,0	**17**

4.1 Architecture, training and validation of Bayesian Regularization

Figure 3 illustrates the winning configuration for the RNA model with the Bayesian Regularization training algorithm.

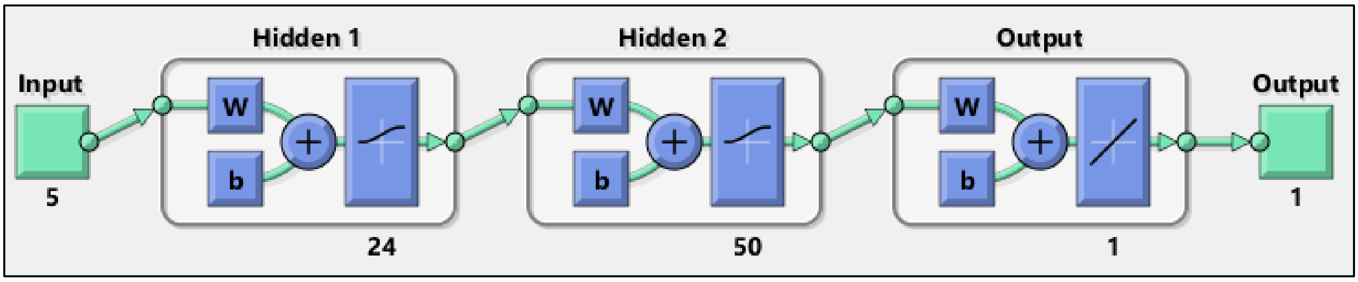

Figure 3: Winning RNA architecture.

It is worth noting that the input layer has 5 neurons referring to the input variables and 1 output neuron which refers to the Reliability forecast result. Table 6 shows the values obtained using the best RNA selection algorithm.

Table 6: Result of the best ANN using Bayesian Regularization.

FC1	FC2	NC1	NC2	Épocas	MSE	RMSE	MAPE
Sigmoide	Sigmoide	24	50	24	0.0000000 001	0.0000104 202	0.0000372 952

According to the results of tests carried out, the winning network model obtained a rate of 0.0000104202 of RMSE and a configuration of 5 neurons in the input layer, 24 in the first intermediate layer, 50 in the second intermediate layer and 1 in the output layer, the it even reaches its state of convergence at the time 24 of 1000 using the Sigmoide transfer function in the input and intermediate layers. Figure 4 illustrates the best training performance of the network where the mean quadratic error obtained from the training step is 0.0000104202 or 1.029e-10.

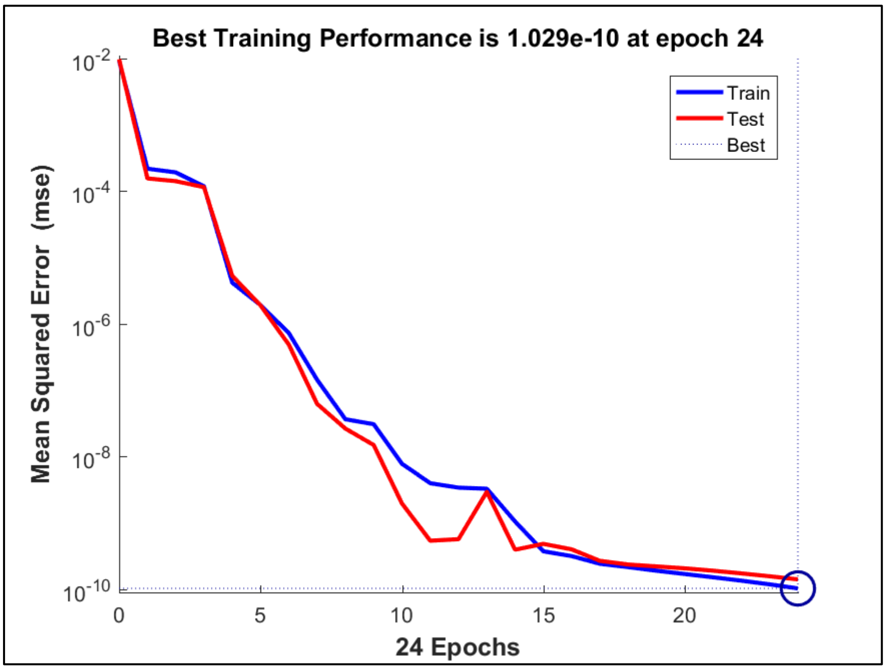

Figure 4: Graph of the best training performance.

To prove the effectiveness of the network through a statistical error analysis, figure x illustrates the error results from the smallest to the largest obtained through the training, test and validation steps, where the smallest error is achieved in the Zero Error marking line.

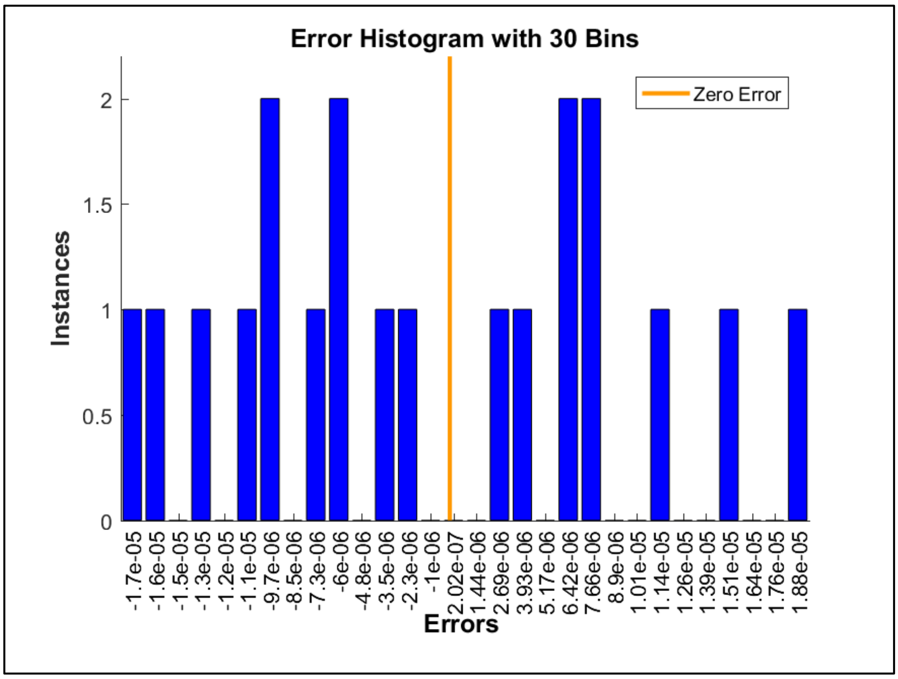

Figure 5: Histogram of error performance.

4.2 Network simulation applied to the engine group Reliability forecast

The simulation was performed by the numerical calculation software known as MatLab 2016a where it was possible to simulate the real data and then compare it with the simulated data from the network. Table 7 presents the results obtained from the calculated model.

Table 7: Calculated Model Results.

Motor	Calculated reliability %	Calculated failure %
1	24,70140556	75,29859444
2	24,72584755	75,27415245
3	24,69794165	75,30205835
4	24,72949717	75,27050283
5	24,68052799	75,31947201
6	24,72562121	75,27437879
7	24,70145604	75,29854396
8	24,68055314	75,31944686
9	24,70488315	75,29511685
10	24,71875172	75,28124828
11	24,71877582	75,28122418
12	24,71193255	75,28806745
13	24,71884821	75,28115179
14	24,71544435	75,28455565
15	24,70499279	75,29500721

16	24,71188414	75,28811586
17	24,72579987	75,27420013
18	24,69449831	75,30550169
19	24,70138317	75,29861683
20	24,72582587	75,27417413

The simulation was carried out with the previously validated configurations of the Bayesian Regularization algorithm to predict the probability of operation of the motor group in a time span of 22 days. Table 8 presents the relative data of predicted reliability and predicted failure.

Table 8: Results of the Predicted Model.

Motor	Expected reliability %	Predicted failure %
1	24,70140556	75,29859444
2	24,72584755	75,27415245
3	24,69794165	75,30205835
4	24,72949717	75,27050283
5	24,68052799	75,31947201
6	24,72562121	75,27437879
7	24,70145604	75,29854396
8	24,68055314	75,31944686
9	24,70488315	75,29511685
10	24,71875172	75,28124828
11	24,71877582	75,28122418
12	24,71193255	75,28806745
13	24,71884821	75,28115179
14	24,71544435	75,28455565
15	24,70499279	75,29500721
16	24,71188414	75,28811586
17	24,72579987	75,27420013
18	24,69449831	75,30550169
19	24,70138317	75,29861683
20	24,72582587	75,27417413

According to figure 6, it is possible to compare the calculated and predicted models where the correlation between them is given by 0.0000104202 of mean square error and 99.9% correlation. According to the results obtained from the predicted model, the engine with the highest failure rate is number 4 with 75.28% and the engine with the lowest failure rate is number 5 with a rate of 75.32 %.

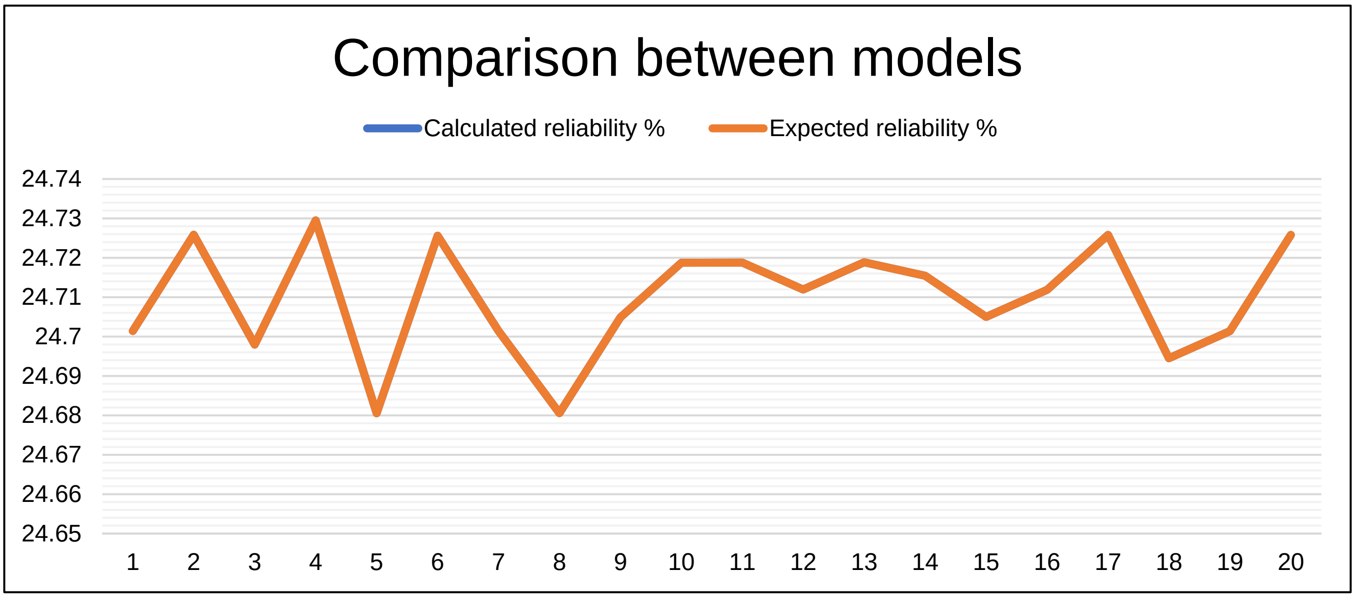

Figure 6: Comparative chart between models.

5. Conclusion

During the research it was possible to identify new research possibilities, considering that in addition to the variables raised (Total Downtime, Total Frequency, Total Occurrences, Average Time Between Failures, Failure Rate, Total Days and Reliability) for the model could be added to increase the consistency and accuracy of the network.

With this, new methodologies can be applied as is the case with Self-Organizing Maps (SOM) for classification of patterns, in this way, it would be possible to determine fault characteristics and determine the probability of new events, or even the use of the supervised approach. considering other fault identification characteristics.

Among the models of RNA architectures applied for prediction and analysis of the reliability KPI, the Bayesian Regularizers with the configurations of 5 neurons in the input layer, 24 in the first hidden layer, proved to be able and accurate to work with the proposed data. 50 in the second hidden layer and 1 in the output layer, presenting a rate of 0.0000104202 of RMSE, accurately estimating the reliability of the motor group. When comparing the results between the calculated and predicted models, it is possible to identify the similarity in the 22-day projection, according to the predicted model, it was possible to achieve a 99.9% hit rate.

6. Acknowledgement

Acknowledgments to the Institute of Technology and Education Galileo da Amazônia - ITEGAM and to the Graduate Program in Engineering, Process Management, Systems and Environmental - PPG-EGPSA.

7. References

ABDULRAHMAN, Shaymaa Adnan et al. Comparative study for 8 computational intelligence algorithms for human identification. **Computer Science Review**, v. 36, p. 100237, 2020.

AGNESE, Marco Antônio Dall. Análise da confiabilidade da manutenção em tratores de uma empresa de produção agrícola. 2020.

ARABI BULAGHI, Zohre et al. World competitive contest-based artificial neural network: A new class-specific method for classification of clinical and biological datasets. 2020.

ARUNTHAVANATHAN, Rajeevan et al. Fault detection and diagnosis in process system using artificial intelligence-based cognitive technique. **Computers & Chemical Engineering**, v. 134, p. 106697, 2020.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS (ABNT). NBR 5462: confiabilidade e mantenabilidade - terminologia. Rio de Janeiro, 1994.

AYVAZ, Serkan; ALPAY, Koray. Predictive maintenance system for production lines in manufacturing: A machine learning approach using IoT data in real-time. **Expert Systems with Applications**, v. 173, p. 114598, 2021.

BAI, Bin et al. Fault data screening and failure rate prediction framework-based bathtub curve on industrial robots. **Industrial Robot: the international journal of robotics research and application**, 2020.

BARBOSA, Douglas AM; FERREIRA, Vitor H. Inferência Bayesiana Aplicada a MLPs para Previsão Probabilística de Carga Semanal. **Simpósio Brasileiro de Sistemas Elétricos-SBSE**, v. 1, n. 1, 2020.

BAROROH, Dawi Karomati; CHU, Chih-Hsing; WANG, Lihui. Systematic literature review on augmented reality in smart manufacturing: Collaboration between human and computational intelligence. **Journal of Manufacturing Systems**, 2020.

CABEZA, R. Torres et al. Faults Diagnostic using Hopfield Artificial Neural Network in front of Incomplete Data. **Journal of Engineering and Technology for Industrial Applications-JETIA**, v. 4, n. 13, p. 6, 2018.

CARDOSO, Diogo Emanuel Da Rocha. Aplicação de conceitos de manutenção preditiva com aplicação de ferramentas de Inteligência Artificial. 2020.

CHEN, Xiang et al. On the role of crack tip creep deformation in hot compressive dwell fatigue crack growth acceleration in aluminum and nickel engine alloys. **International Journal of Fatigue**, v. 145, p. 106082, 2021.

CHINI, Christopher M.; LOGAN, Lauren H.; STILLWELL, Ashlynn S. Grey water footprints of US thermoelectric power plants from 2010–2016**. Advances in Water Resources**, v. 145, p. 103733, 2020.

CORRÊA, Rafaela Gomide. Estudo numérico do escoamento de ar em um motor de combustão interna. 2020.

EL-ADAWY, Mohammed et al. Stereoscopic particle image velocimetry for engine flow measurements: Principles and applications. **Alexandria Engineering Journal**, v. 60, n. 3, p. 3327-3344, 2021.

FERREIRA, Vitor Hugo; DE SOUZA, Julio Cesar Stacchini; DO COUTTO FILHO, Milton Brown. Inferência Bayesiana Aplicada ao Desenvolvimento de Modelos Neurais para Tratamento de Alarmes em Subestações, 2020.

FONSECA-JUNIOR, M. et al. Programa de gestión de mantenimiento a través de la implementación de herramientas predictivas y de TPM como contribución a la mejora de la eficiencia energética en plantas termoeléctricas. **Dyna**, v. 82, n. 194, p. 139-149, 2015.

HAN, Xiao et al. Remaining useful life prediction and predictive maintenance strategies for multi-state manufacturing systems considering functional dependence. **Reliability Engineering & System Safety**, v. 210, p. 107560, 2021.

Koçak, Y., & Üstündağ Şiray, G. New activation functions for single layer feedforward neural network. **Expert Systems with Applications**, 164, 113977. doi:10.1016/j.eswa.2020.113977. 2021.

LEOCÁDIO, Caio Monteiro; FERREIRA, Vitor Hugo. Inferência Bayesiana no desenvolvimento de previsores neurais de vazão diária utilizando informações de precipitação. **Journal of the Brazilian Neural Network Society**, v. 10, n. 3, p. 157-165, 2012.

LU, Xue-Qin et al. Metaheuristics for homogeneous and heterogeneous machine utilization planning under reliability-centered maintenance. **Computers & Industrial Engineering**, v. 151, p. 106934, 2021.

LUGHOFER, Edwin; SAYED-MOUCHAWEH, Moamar. Predictive maintenance in dynamic systems: advanced methods, decision support tools and real-world applications. **Springer**, 2019.

MANHERTZ, Gabor; BERECZKY, Akos. STFT spectrogram based hybrid evaluation method for rotating machine transient vibration analysis. **Mechanical Systems and Signal Processing**, v. 154, p. 107583, 2021.

MEIßNER, Christian et al. Investigation on wall and gas temperatures inside a swirled oxy-fuel combustion chamber using thermographic phosphors, O2 rotational and vibrational CARS. **Fuel**, v. 289, p. 119787, 2021.

MORO, Giancarlo Dal. Efficient Joint Analysis of Surface Waves and Introduction to Vibration Analysis: Beyond the Clichés. **Springer Nature**, 2020.

RIGHETTO, Sophia Boing et al. Manutenção Preditiva 4.0: Conceito, Arquitetura e Estratégias de Implementação. 2020.

ROCHA, Márcio Andrade et al. Aplicação da análise de vibração na determinação do atraso de ignição em um motor de combustão interna por compressão. **Brazilian Journal of Development**, v. 6, n. 12, p. 99947-99952, 2020.

RUIZ-HERNÁNDEZ, Diego; PINAR-PÉREZ, Jesús M.; DELGADO-GÓMEZ, David. Multi-machine preventive maintenance scheduling with imperfect interventions: A restless bandit approach. **Computers & Operations Research**, v. 119, p. 104927, 2020.

SALLES, Gisele Maria de Oliveira et al. Estimação de intervalos de tempo ótimos para a inspeção e manutenção de escovas em unidades geradoras da copel. 2020. Dissertação de Mestrado. **Universidade Tecnológica Federal do Paraná**.

SÁNCHEZ, D. et al. Experimental enhancement of a CO2 transcritical refrigerating plant including thermoelectric subcooling. **International Journal of Refrigeration**, v. 120, p. 178-187, 2020.

SCHWENDEMANN, Sebastian; AMJAD, Zubair; SIKORA, Axel. A survey of machine-learning techniques for condition monitoring and predictive maintenance of bearings in grinding machines. **Computers in Industry**, v. 125, p. 103380, 2021.

SILVA, Jean da Silva de Abreu et al. PROPOSTA DE IMPLANTAÇÃO DE SISTEMA DE PROTEÇÃO CONTRA POTENCIAL DE FALHA DO MOTOR À DIESEL (DISPARO DO MOTOR). **ITEGAM-JETIA**, v. 5, n. 19, p. 06-11, 2019.

SOLTANALI, Hamzeh et al. A comparative study of statistical and soft computing techniques for reliability prediction of automotive manufacturing. **Applied Soft Computing**, v. 98, p. 106738, 2021.

SOUZA, Arthur Gabriel. REDESIGN DE MÁQUINA EMBALADORA TERMOENCOLHIVEL COM BASE NA METODOLOGIA TOOLBOX PARA INDUSTRIA 4.0. **Engenharia de Produção-Pedra Branca**, 2020.

TIAN, Hua; LIU, Peng; SHU, Gequn. Challenges and opportunities of Rankine cycle for waste heat recovery from internal combustion engine. **Progress in Energy and Combustion Science**, v. 84, p. 100906, 2021.

YU, Xinnan; FENG, Zhipeng; LIANG, Ming. Analytical vibration signal model and signature analysis in resonance region for planetary gearbox fault diagnosis. **Journal of Sound and Vibration**, v. 498, p. 115962, 202.

ZINGONI, Alphose. Use of symmetry groups for generation of complex space grids and group-theoretic vibration analysis of triple-layer grids. **Engineering Structures**, v. 223, p. 111177, 2020.

ZOU, Guang et al. Fatigue inspection and maintenance optimization: A comparison of information value, life cycle cost and reliability based approaches. **Ocean Engineering**, v. 220, p. 108286, 2021.