International Journal of Computer Information Systems and Industrial Management Applications.
ISSN 2150-7988 Volume 15 (2023) pp. 531-540

Received: 06 September 2022; Accepted: 15 May 2023; Published: 9 July 2023

Application of Data Association and Perceptron Artificial Neural Networks (AR-ANN) in Fault Detection in Dynamic Systems: Gears

Paulo Roberto Tebon[1], Roberto Outa[2], Fábio Roberto Chavarette[3], Aparecido Carlos Gonçalves[4], Sandro da Silva Pinto[5], Samuel Stabile[6]

[1] Araçatuba College of Technology,
Av. Prestes Maia, 1764 - Ipanema, Araçatuba - SP, 16075-051, Brasil,
paulo.tebon@fatec.sp.gov.br

[2] Lins College of Technology,
Estrada Mário Covas Junior , km 1 - Vila Guararapes, Lins - SP, 16403-025, Brasil,
roberto.outa@fatec.sp.gov.br

[3] UNESP - Paulista State University - Department of Engineering, Physics and Mathematics of the Institute of Chemistry
Av. Prof. Francisco Degni, 55 - Jardim Quitandinha, Araraquara - SP, 14800-900, Brasil
fabio.chavarette@unesp.br

[4] UNESP - Paulista State University - Faculty of Engineering of Ilha Solteira, Department of Mechanical Engineering
Avenida Brasil, 56 - Centro - Ilha Solteira/SP - CEP 15385-000, Brasil
aparecido.carlos@unesp.br

[5] Lins College of Technology
Estrada Mário Covas Junior , km 1 - Vila Guararapes, Lins - SP, 16403-025, Brasil
sandro.pinto@fateclins.edu.br

[6] Lins College of Technology
Estrada Mário Covas Junior , km 1 - Vila Guararapes, Lins - SP, 16403-025, Brasil
samuel.stabile@fatec.sp.gov.br

***Abstract*: This work demonstrates a study of identification, classification and grouping of different signals, whose objective is the detection of failures between a pair of gears. Therefore, it is a multidisciplinary work, as it promotes an application of low-cost embedded systems and methodologies of computer science in the area of mechanical engineering. For this to be done, the concept of perceptron artificial neural networks (ANN) associated with the data association rules (AR) theorem belonging to the concept of data-mining was used. This association was developed because it is easy to access and has great potential in identification and classification. We named these different theorems AR-ANN. The result of the application of AR-ANN to the reference and faulty signs was successful, whose classification demonstrated a high rate of correct and in the training phase of the perceptron network, the balance of the adjustment line was obtained, demonstrated by linear regression and weights (variables).**

Keywords: Artificial Neural Network-ANN; Data-Mining; Vibration; Bioengineering; Fault Detection; Association Rules-AR**.**

I. Introduction

Over the years we have observed several changes in industries and especially the evolution of different technologies, generating new forms of work with greater efficiency. The adoption of new technologies in industries is considered industry 4.0, which uses cyberphysical systems, internet of things, cloud computing, robotics, artificial intelligence, biotechnology, among others [1].

Industrial systems require a greater effort of the management system to keep machinery and equipment in a good working state. Generally, equipment with insufficient repairs and proper maintenance is more likely to fail and damage. The result is low productivity and increased costs inherent in the process [2]. Failure identification methods can determine the severity level of a failure, which is made by specific software that uses methods linked to artificial intelligence, deep learning, artificial neural network, among others [3].

Several researchers have developed activities related to artificial intelligence methods, in one of these studies the neural network approach was developed to detect and identify the failures of a flight control system, whose problem is solved with the introduction of estimators of neural networks of online learning [4]. Another researcher uses the concept of artificial neural networks and deep learning for the detection and classification of failures, whose problem was solved by classifying the failures, later, through network training, it was possible to detect the failures considering also the hidden layers [5]. Another author studies the behavior of a structure in order to obtain information of damage to the structure and, for this, uses intelligent computing technique linked to SHM. The result shows that the proposed methodology is robust and accurate [6]. Another author uses the concept of artificial immune systems (AIS) applied to structural health monitoring to detect failures in a dynamic rotor, whose problem was solved with concepts of monitoring and sensing of the AIS, classifying and grouping the signals in the degree of severity of failures and the probability of failures [7].

The objective of this work is to propose a new multidisciplinary method with application in the determination of failures, using the Wiener filter, artificial neural networks (ANN) and association rules (AR). For this to be done, this work was divided at first into the decomposition of signals using the Wiener filter, resulting in a database named data-Wiener; then was introduced the concept of separation and classification of grouping using artificial neural network (ANN); finally, the concept of data association (AR) was applied.

II. Gear: Dynamic Elements

The gears are intended to transmit the movement from one rotary shaft to the other, which features such as precise synchronization, efficiency in coupling between gear teeth, gear weight, free slip and uniform movement, are relevant in dynamic system performance. The transmission of force is made considering the contact of two gear teeth until the separation of this pair of gears, whose angular displacement is made by rotation of the shaft [8] [9]. Figure 01 shows the design of two gears considering the moment of coupling between two teeth.

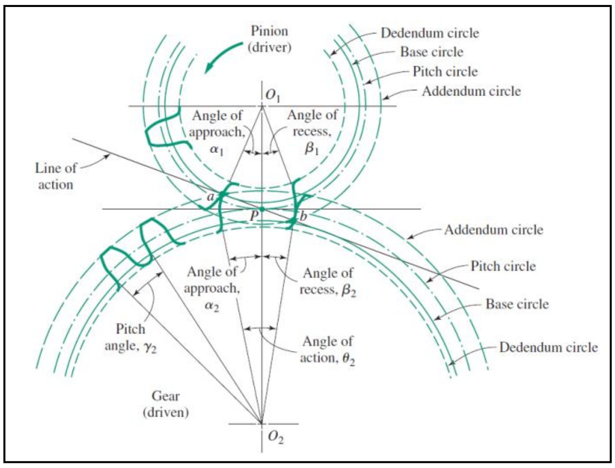


Figure 1. Squematic design of two gears with coupled teeth

The length of action *(Z)* is the distance between the point *a* and *b*, geometrically calculated by the pinion-gear equation,

$$Z = \sqrt{(r_p + a_p)^2 - (r_p cos\varphi)^2} + \sqrt{(r_g + a_g)^2 - (r_g cos\varphi)^2} - Csin\varphi \quad (1)$$

being, r_p the radius of the pinion; a_p is the maximum diameter of the head circle minus the radius of the pinion; r_g is the radius of the gear moved; a_g is the maximum diameter of the head circle minus the radius of the moved gear [8, 9]. Figure 02 shows the interference ratio (contact area) between the spur gear.

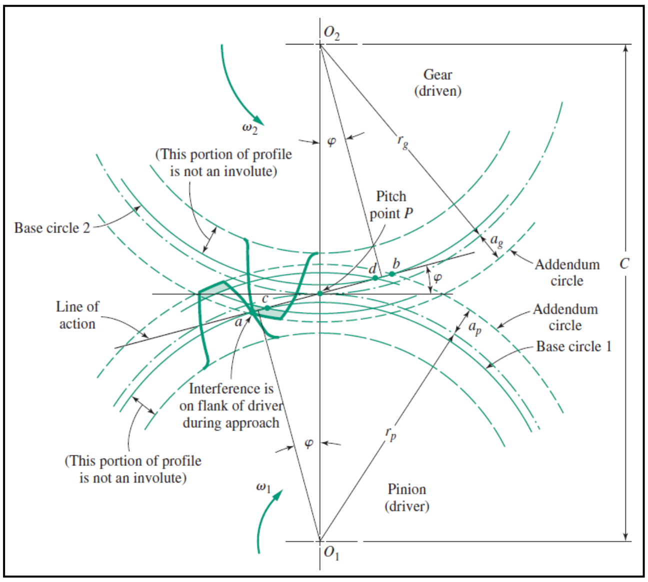


Figure 2. Interference between pinion-gear teeth

Usually, problems in gears are perceived through the observation of a noise or vibration of the system. The problems of the gears are hidden and treatment through the symptoms demonstrated in the system do not arise beneficial effects for an immediate solution of the problem. Therefore, a greater understanding of the system is necessary, considering the problem and possible solutions to these problems [10].

For a gear noise to be generated it is necessary, at first, that there is a variation of force in the system, and the gears when bumping into the contact area, results in the reaction of the force transmitted throughout the structure. The continuous and intermittent shock movement of the gear teeth promote vibration that is distributed throughout the system. The noise generated varies in amplitude, direction and position, as the contact areas move axially along the gear pitch line [10]. Figure 03 shows the schematic drawing of the amplitude, direction and position of the excitation force in the gear.

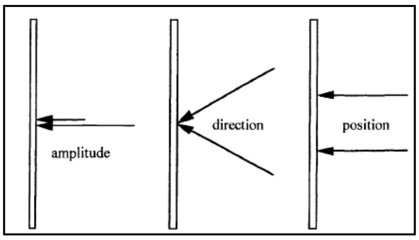


Figure 3. Excitation Force in relation to Amplitude, Direction and Position

When the gear is driven several times, it can position itself in momentarily different locations in relation to the reference position and the speed that should be constant is changed [11].

This demonstrates the transmission error (TE). Technically, the transmission error (TE) is considered as the angular difference between the ideal and actual position of the output gear, resulting in the finite stiffness of gear steaming and its movement errors [9] [12].

In the development of gear designs, the detection of the failure modes of the gear pair is of high complexity, because factors such as the combination of variable loads, with surface slips, among other external sources of action, act directly on the dynamic system [8].

Failures can be seen through the contact tension on the surface of the teeth curves according to the cyclic repetition of the system, with cracks spreading along the gear. Thus, in the contact region, with or without lubricants, abrasive and adhesive wear is analyzed to understand the behavior of the system. Elastic deformation, induced by force in the contact region, is also considered a failure and is directly linked to the stiffness of the gear teeth surface, which significantly contributes to the potentiation of the resonant noise of the gear pair [8] [9] [10].

III. Dynamic Systems and Vibration

In nature different phenomena act in a given body, and, through the function of frequency response (FRF), the behavior of a given dynamic system can be understood. The representation of this system can correspond to the concept of mass-spring-damper, shown in the concepts of vibration [13]. Figure 04 shows a dynamic system composed of mass-spring-damper.

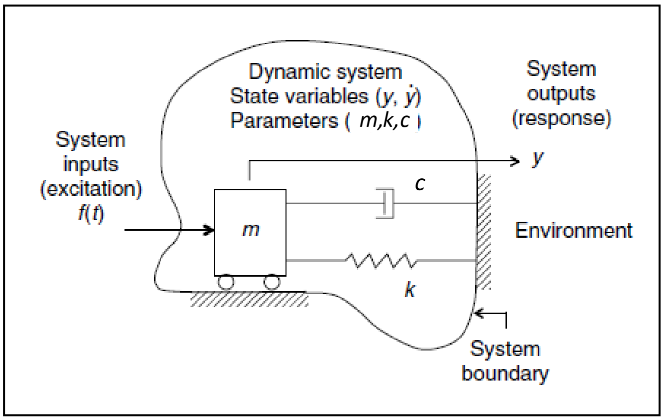


Figure 4. Dynamic System with mass-spring-damper

which, $\ddot{y}$ is acceleration; $\dot{y}$ is the velocity; and the y displacement of the mass m; c represents viscous damping; and stiffness k is a constant [13] [14].

Vibration is linked to the concepts of dynamics and can be understood as the repetition of movements of a given mass in relation to a given reference of nominal position or equilibrium point [15]. A system that oscillates repeatedly and uniformly can be manipulated mathematically and is considered a harmonic motion system [16]. Figure 05 shows the graphical representation of simple harmonic motion.

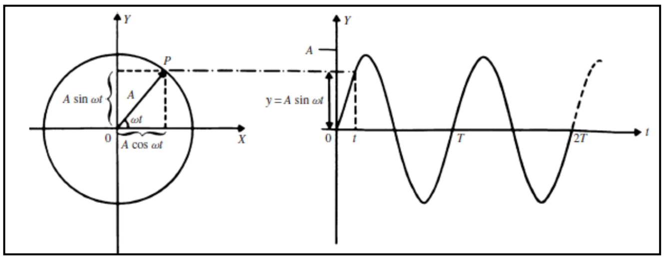


Figure 5. Simple Harmonic Movement

Note that the representation of the P point on the left side of the figure, when in uniform and constant movement counterclockwise, can be located through the geometric concept of trigonometry. When this point P is projected ona two-dimensional cartesian plane, one can see the graphical representation of a sine. The cartesian plane system contains period, frequency and phase information, so the equation representing this system varying in time can be written as:

$$y(t) = Asin(\omega t + \varphi) \tag{2}$$

The argument A corresponds to amplitude and can be represented in the figure as the radius of the circumference; ω is the angular frequency; o t is the time of displacement of the point P; the φ corresponds to the angle of lag of the movement of the point P [16].

In mechanics the degrees of freedoms represent the number of independent movements that a system contains and can be characterized by the concepts of single degree of freedom SDOF, and multi degree of freedom MDOF. Several mechanical systems can be represented through the free body diagram and can be seen in Figure 06 [14] [15] [16].

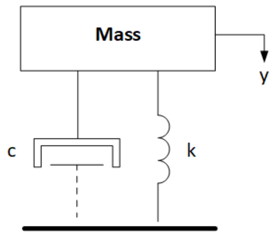


Figure 6. Free Body Diagram

The general equation representing a system in equilibrium can be written as,

$$m\ddot{y} + c\dot{y} + ky = 0 \tag{3}$$

where, $\ddot{y} = \frac{d\dot{y}}{dt}$ is acceleration; $\dot{y} = \frac{dy}{dt}$ is the velocity; and the y displacement of the mass m; c and k are respectively viscous damping and constant stiffness [14] [15] [16].

IV. Digital Signal Analysis

Signal processing is a projected operation that assists in the difference of useful and non-useful information of a given signal. In this condition it is possible to consider that signal processing determines, extraction, enhancement, storage, and transmission of operations for a better condition of the useful signal. This condition can be applied in different areas of engineering, medicine, among others [17] [18]. The different signals can be classified as: Analog and digital signal; Real and complex sign; Periodic and non-periodic signal;

Deterministic and random sign; Single channel and multi-channel signal; Signal of strength and energy; Monocomponent and multicomponent signal [19].

The signs we find in the real world represent information about the behavior of the physical state of a given system under study [18]. In its design they are mainly analog that vary continuously in time, and these signals, when undergoing analog signal processing (ASP) tend to be considered high cost, justified by the lack of flexibility in the processing and complexity of projects. Digital signal processing (DSP) is a more robust and powerful type of processing, as they can be developed using software on general purpose computers. The advantages are based on: component costs are reduced, considering development technology; additions and multiplications operations, the processing of which is carried out at the limit of its capacity, considered stable [17] [19]. The digital signal processing can be shown through a diagram block, shown in Figure 07.

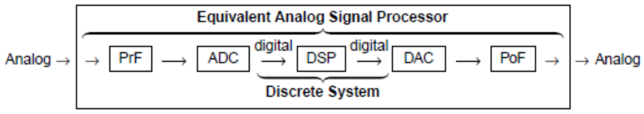


Figure 7. DSP Diagram Block

PrF is an antialiasing filter; ADC is a digital analog converter that produces a flow of binary numbers from analog signals; DPS is a special-use processor, considered a general purpose computer; the DAC converts the digital signal to analog that produces a wave from the sequence of binary numbers; PoF is a powder filter that tends to soften the wave-shaped signal on stairs in the desired analog signal [17].

The modeling of a (T) system, in digital signal analysis, can be defined as a single input signal represents a single output $(x(t))$ signal, and $(y(t))$ mathematically this representation is made in the domain of time. Figure 08 shows the signal processing of a given system under study [19].

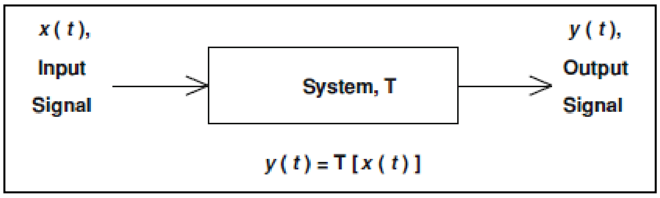


Figure 8. - Signal Processing of a Given System with an Operator

Systems can be classified into: Analog, discrete systems in time and digital; Systems variants in time and not variants in time; Causatous and non-causatis systems; Static and dynamic systems; Stable and non-stable systems; and finally, linear and non-linear systems. Thus, for each type of system a specific mathematical concept is introduced to understand its behavior [18] [19].

Time domain systems or time functions use the Fourier transform principle, which decomposes a temporal function into a complex value function in the frequency domain. The result is an absolute value that represents the sum of the frequencies present in the original function, and the complex argument determines the displacement phase of that frequency [18].

The cross-correlation method is also used in the analysis of signals, whose purpose is to understand the behavior between two signs, the result of $(\sigma_x;\ \sigma_y)$ which can be obtained considering the correlation coefficient, r applying the least squares theorem or curve adjustment. The equation of the correlation coefficient r, can be written as [20],

$$r = \left[1 - \frac{\sigma_{y,x}^2}{\sigma_y^2}\right]^{1/2} \tag{4}$$

Another theorem that can be applied to domain systems in time is the Wiener filter, which produces an estimate of a random process desired by invariant linear filtering at the time of a given signal. The Wiener filter minimizes the average quadratic error between the estimated random process and the desired process and is applied to eliminate unwanted noise from a given signal [19] [21]. The equation can be written as,

$$H_0(s) = \frac{S_{yx}(s)}{S_{xx}(s)} e^{\alpha x} \tag{5}$$

where α it represents the Wiener filter delay $H_0(s)$; $S_{xx}(s)$ represents signal power spectrum (PSD) of the original signal; the $S_{yx}(s)$ represents the crossspectral density (CSD) between the original and observed signals [19].Considering the spectral formulation presented in equation 5, the equation representing the Wiener (W_f) filter can be rewritten based on minimizing the mean quadratic error between the observed signal and a reference signal, which is equal to the original signal,

$$W_f = \left[R_{yy}^{-1} R_{yd}\right]^T \tag{6}$$

where, R_y is the observed signal; R_{yy} is the autocorrelation of the observed signal; and R_{yd} it is the cross-correlation of the signal [19].

The signals of a given physical model can be reproduced in experimental projects, whose main idea is to understand behavior using experimental methods, unlike only analytical conditions commonly treated in engineering. The purpose is to obtain accurate quality results, knowing the measurement estimates [22].

V. Data Mining and Data Association

The comparison of methods for classifying objects into groups can be done based on a statistical approach, or on the computer science approach, whose goal is to develop a data set model that assigns future objects into predetermined classes. The reliability of the methods is done by cross-validation or by applying a model to a new test set of similar objects, validating the error rates. Discriminant analysis determines the estimation of the distance from the object to the center of the class and assigns the object to the nearest group [23].

A Data Mining

Currently several technologies are applied for data collection, being continuous or periodic basis, whose data allow the understanding of the physical behaviors of the system under study in several areas. However, for the analysis of a large amount of data (large data) it is necessary to introduce automatic data mining tools. The objective of classification and regression is to build a model that can be used as a prediction. In the classification the construction of a concise model of the distribution of the attribute dependent on the attributes of the predictor, which results in the model of known and unknown predictor attribute values of the dependent attribute. This application can be developed in scientific experiments [24][23].

The association rules use two distinct concepts, the first is the rule of classification by induction, and the second is the rule of association by discovery, and both use rules to characterize regularities in a data set [23][25]. In particular, the induction classification rule is intended to acquire the ability to make predictions, while the discovery membership rule provides understanding (insight) to the user. The application of each rule for each situation in the membership rules is determined according to the user's value, applied in the representation of the data to be analyzed [23].

The discovery association rule determines the association (affinity) between two sets, considered antecedent and consequential, which the result can contain a single item to be associated [25]. Considering that the left side is the antecedent *or left-hand side* (LHS), and the right side is the *right-hand side* (RHS) is the consequent, the affinity between the two sets is made through statistics based on the frequency of relationship [23].

The discovery association rule theorem determines that "it is D the transaction database and the number of and N transactions in D. Each transaction D_i is a set of items. Be support the (X) proportion of transactions that contain the set of items X" [23]:

$$Support(X) = \frac{|\{I | I \in D \wedge \sqsupseteq X\}|}{N} \quad (7)$$

"where I it is a set of items and $|\cdot|$ demonstrates the cardinality of a set. The support in the membership rule is the proportion of *transactions* that the antecedent and the consequential contain. With this, the confidence of *the* binding rule is the proportion of the transaction, that is $A \rightarrow C$" [23].

$$support\ (A \rightarrow C) = support\ (A \cup C) \quad (8)$$

$$confidence\ (A \rightarrow C) = \frac{support(A \cup C)}{support\ (A)} \quad (9)$$

"If *the support* is high enough, then confidence is a reasonable estimate of the probability of any future transaction containing the antecedent and containing the consequent one" [23]. The pseudocode that demonstrates the join rule can be seen in Figure 9.

```
L_1 = {frequent one-item-item sets}
for k = 2; L_{k-1} ≠ Ø; k++ do begin
    C_k = {{x_1, x_2, ..., x_{k-2}, x_{k-1}, x_k} | {x_1, x_2, ..., x_{k-2}, x_{k-1}} ∈ L_{k-1} ∧
          {x_1, x_2, ..., x_{k-2}, x_k} ∈ L_{k-1}}
    for all_transactions t ∈ D do begin
        for all candidates c ∈ C_k ∧ c ⊆ t do
            c.count++;
    end
    Lk = {c ∈ C_k | c.count ≥ minsup}
end
return ⋃_k L_k;
```

Figure 9. Pseudocode of the Association Rule by Discovery

B Data Mining and Artificial Neural Networks

Neural networks are the main technology used for data mining, and this technology began with McCullough and Pitts in 1943 with an article on binary neuron using the human brain as a computational model. Around 1960 digital computing initiates numerical manipulation to represent symbols, that is, the transformation of symbolic abstraction into numbers by computational processes. With this, different structural models are developed as the perceptron by Frank Rosenbaltt, and the Adaline by Widrow and Hoff. Perceptron was conceived using the mental model and reserved for 32-bit operating systems, while Adaline solves difficult engineering problems such as adaptive filtering [26].

Perceptron is an interactive procedure that adjusts the increment w and θ, until the decision limit is able to separate into two training classes. In this case it is possible to demonstrate the effect of a data measurement noise by separating them into classes. When an external source is introduced into the data measurement noise it is necessary to introduce a technique called support vector machines (SVM) which is a supervised learning method that analyzes data and recognizes patterns, that is, receives new data at the input and generates a predicted data output [25] [27]. The SVM containing a sigmoid kernel is equivalent to a simple two-layer neural network, known as multilayer perceptron (without hidden layers) [25].

The perceptron network belongs to the single-layer feedforward architecture, whereas the flow of information flows from the input layer to the output layer, without any feedback produced by the neuron. Thus, the perceptron network, consisting of n input signals and only one output, can be seen in Figure 10 [28].

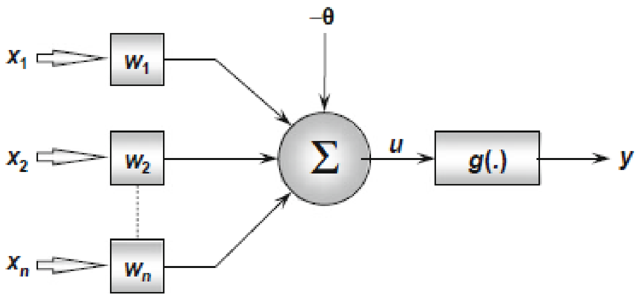


Figure 10. Perceptron network of a single output neuron

Thus, $X_1, \ldots, X_n$ the input variables of a process correspond; $w_1, \ldots, w_n$ correspond to the weights (weightings); the θ is the activation threshold; the activation u potential; the is the activation $g(\cdot)$ function; the output $y$corresponds [28]. The equation corresponding to perceptron, can be written as,

$$u = \sum_{i=1}^{n} w_i X_i - \theta \quad (10)$$

$$y = g(u) \quad (11)$$

the signal activation function is written as [28],

$$y = \begin{cases} 1,\ se \sum w_i X_i - \theta \geq 0 \leftrightarrow w_1 X_1 + w_2 X_2 - \theta \geq 0 \\ -1,\ se \sum w_i X_i - \theta < 0 \leftrightarrow w_1 X_1 + w_2 X_2 - \theta < 0 \end{cases} \quad (12)$$

The perceptron training process, based on Hebb's learning rule (1949) determines that the output produced by the perceptron is not coincident with the desired output, the sinatic weights and activation thresholds of the network should be increased in proportion to the values of its input signals. This process is done repeatedly and sequentially in all training samples, until the perceptron output is similar to the desired output [26] [28] [29]. The rules of adjustments and weights can be defined as,

$$\begin{cases} w_i^{atual} = w_i^{anterior} + \eta(d^K - y)X_i^K \\ \theta_i^{atual} = \theta_i^{anterior} + \eta(d^K - y)(-1) \end{cases} \tag{13}$$

simplifying the equation, we will have,

$$w^{atual} = w^{anterior} + \eta(d^K - y)X^K \tag{14}$$

where, η is the learning parameter (step), is the training K sample, and d is the desired value for the K-th training sample [26] [28].

The optimization problem consists of an objective function that one wants to discover the maximum and minimum values over a set of constraints that satisfy the best value of the objective function [26] [27]. Thus, different optimization techniques can be applied as:

- descending gradient with error backpropagation;
- descending gradient with adaptive learning rate backpropagation;
- descending gradient with impulse and backpropagation of the adaptive learning rate;
- almost Newton backpropagation of Broyden-Fletcher-Goldfarb-Shanno (BFGS);
- backpropagation of Bayesian regularization;
- conjugated gradient of backpropagation with powell-beale restarts,
- conjugated gradient of backpropagation with fletcher-reeves updates;
- gradient conjugated with backpropagation with Polak–Ribiere updates;
- gradient conjugated with scale dwell;
- Levenberg-Marquardt algorithm;
- secant retropropagation in one step [24].

The training pseudocode can be shown in Figure 11 [28],

```
Begin {PERCEPTRON Algorithm – Training Phase}
  <1> Obtain the set of training samples {x^(k)};
  <2> Associate each desired output {d^(k)} to each sample;
  <3> Initialize vector w with small random values;
  <4> Specify the learning rate {η};
  <5> Initialize the epoch counter {epoch ← 0};
  <6> Repeat the following instructions:
        <6.1> error ← "none";
        <6.2> For all training samples {x^(k), d^(k)}, do:
                <6.2.1> u ← w^T · x^(k);
                <6.2.2> y ← signal(u);
                <6.2.3> If y ≠ d^(k)
                      <6.2.3.1> then { w ← w + η · (d^(k) − y) · x^(k)
                                     { error ← "existent"
        <6.3> epoch ← epoch + 1;
     Until: error ← "none"
End {PERCEPTRON Algorithm – Training Phase}
```

Figura 11. Training Pseudocode

The pseudooperation code can be shown in Figure 12 [28],

```
Begin {PERCEPTRON Algorithm – Operation Phase}
  <1> Obtain one sample to be classified {x};
  <2> Use the vector w adjusted during training;
  <3> Execute the following instructions:
        <3.1> u ← w^T · x;
        <3.2> y ← signal(u);
        <3.3> If y = −1
              <3.3.1> then: sample x ∈ {Class A}
        <3.4> If y = 1
              <3.4.1> then: sample x ∈ {Class B}
End {PERCEPTRON Algorithm – Operation Phase}
```

Figure 12. Operation Pseudocode

VI. Experimental Methodology

The experiment representing the dynamic system with gears is based on a simple system of a rotor coupled by pulleys to the main power emission source, an electric motor. Coupled to this rotor, on the main shaft, is the main gear, which is coupled to another gear supported on a bracket. Figure 13 shows the coupled gear rotor assembly.

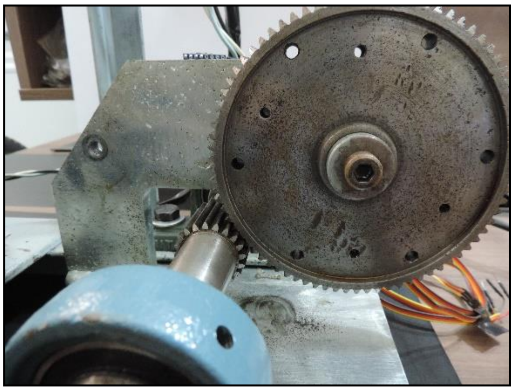


Figura 13. Rotor with Coupled Gears.

The figure shows the main shaft coupled to a gear, but it supports the ability to fasten to two gears, one on the left side and the other on the right side. In this experiment, only a single gear will be used for fault detection, which is on the right side.

Basically, the main shaft movement is done by an electric motor of 0.5hp – 60Hz that has the speed of the motor shaft altered by a CFW 10 – WEG frequency inverter that acts from 0.1Hz to 60Hz. The main shaft is 1045 steel with a diameter of 25.0 mm, the main shaft gear has a maximum tooth height close to 23.0mm, and the coupling gears have the maximum height of the teeth close to 123.0 mm. The bearing bearings, supporting the main shaft, are both P205 UC205.

The experiment starts in the 4Hz setup of the frequency inverter and given the power transmission of the motor to the main axis of the gears, the movement of the main shaft is transmitted the gear. The variation of the main axis rotation occurs every 2Hz, whose maximum frequency value is 14Hz. For the collection of the vibration signal, the MPU6050 sensor was used, which were positioned one in the housing and the other in the support of the gear fixation structure. Figure 14 shows the positioning of the MPU6050 sensors.

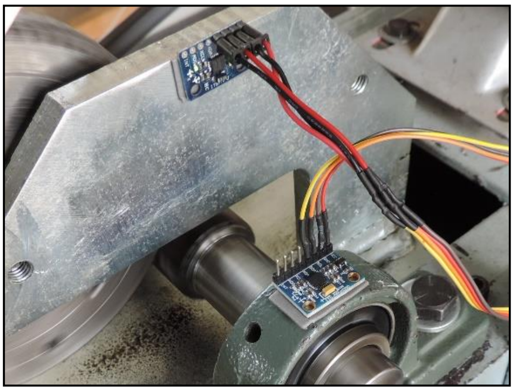


Figure 14. Positioning the MPU6050 Sensors

The figure shows the position of the sensors, whose purpose is to capture the signal of the structure and housing for further analysis.

The signal sample is 4000 samples and for sample validation, the finite sample theorem was used. Through the theorem the minimum amount of samples would be 1111, considering a sampling error of 2% and confidence level of 95%. Therefore, it is considered that the 4000 samples can be used.

The formation of the database follows the following characteristic: the first is the reference vibration signal; the second is the faulty signal of the dynamic system. The faulty signal composition of the dynamic system is divided by the vibration collection in the gear fixing bracket, and in the bearing housing, as shown in Figure 14. Each signal contains a size of 1x 4000, and data collection starts at 4Hz at 14Hz, varying every 2Hz, controlled by the frequency inverter. Five signals of each frequency were collected. The fault condition is determined by the characteristic of the gear coupling to the bracket, because in the bracket, the gear fixing hole contains a purposeful clearance, thus allowing the unbalance of the centering of the coupling gear.

The result of the database of this experiment has the reference signal matrix of size 30 x 4000, and the faulty signal matrix is 60 x 4000 in size.

From which the samples were collected, these samples were processed by the Wiener filter, whose idea is to eliminate unwanted noises. Figure 15 shows the application of the Wiener filter on the 4Hz vibration signal.

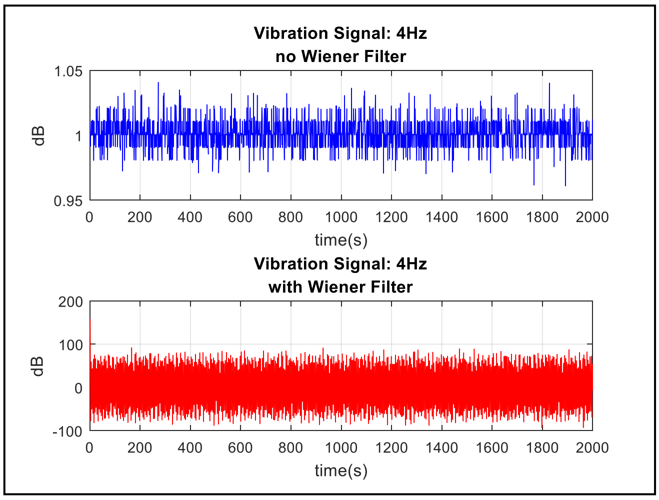


Figure 15. 4Hz Vibration Signal with Wiener Filter

Note that the Wiener filter has demonstrated its applicability when it acts to minimize the average quadratic error, and that it can be seen in signal amplitude (dB). Therefore, consider that all collected signals of vibration of the housing and gear support, were submitted to wiener filter application. The linear regression of the 4Hz vibration signal can be shown in Figure 16.

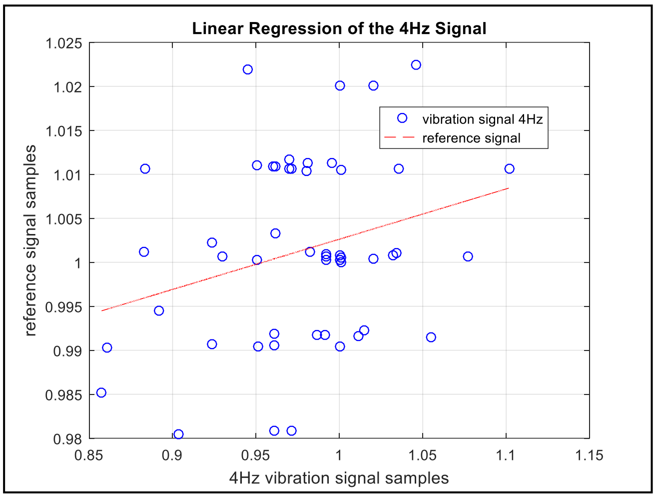


Figure 16. Linear Regression of the 4Hz Vibration Signal

It is noted that the 4Hz vibration signal, coupled to the main axis, demonstrates a different shape from the reference signal, because thus, the values of the variables are far from the best fit line. The residual error is 0.09332, $(r = 0{,}09332)$. The line equação is $0{,}057126x + 0.945502$. This result demonstrates and fortifies the condition of the existence of the difference between the reference signals and the faulty signals.

The database, now called data-Wiener, is processed by the neural networks algorithm containing the data-mining association rule theorem. Thus, this working model can be called *artificial neural network – association rules* (AR-ANN). AR-ANN was processed throughout the data-Wiener database, considering that the parameter with the reference characteristic is the reference signal. Figure 17 shows the result of the cross-reference test.

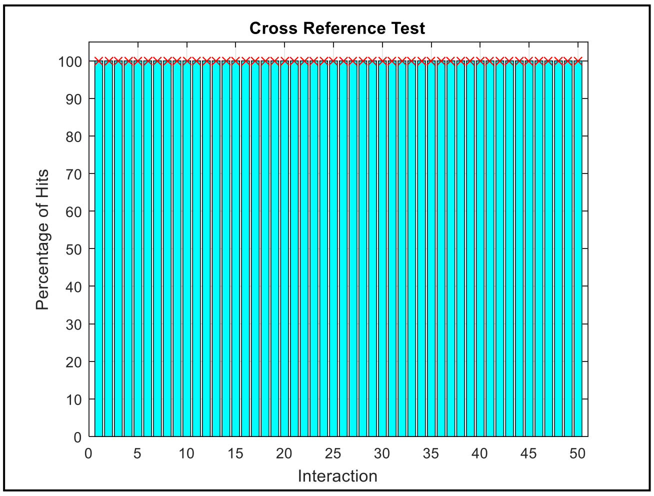


Figure 17. Cross-Reference Test

Figure 18 shows the result of linear regression analysis.

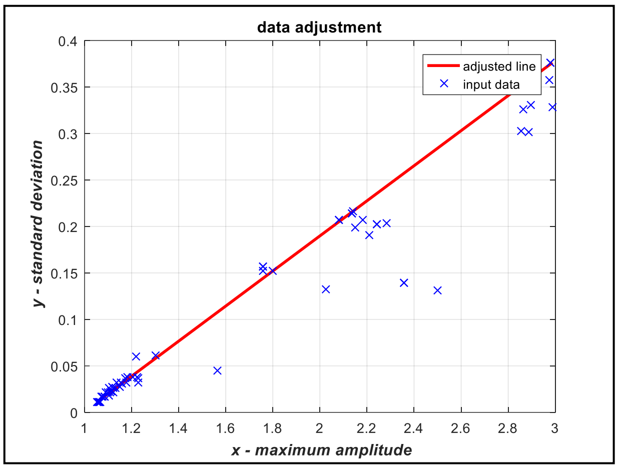


Figure 18. Linear Regression Analysis

Figure 19 shows the learning curve of database interaction

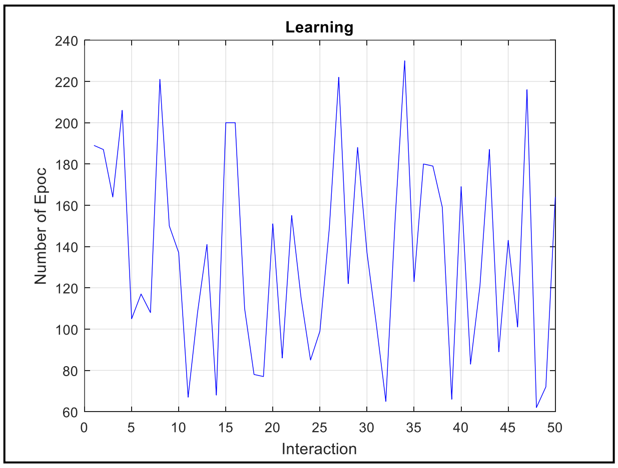


Figure 19. Learning curve

Note that the network has been trained with 164 times. The season corresponds to a forward passage and a passage backwards. The result for this amount of season defines 100% hit (cross-reference) and the weights are 0.0027; 0.0614 and 0.4963. Specifically each weight corresponds to the value of a vector or neuron.

Table 1 shows the results of the metrics obtained from the application of the AR-ANN method.

Table 1. AR-ANN Metric Result

Metrics	Description	%
Reliability	Reliability explains the probability of the relationship between the data (data-Wiener)	100
Accuracy	Accuracy explains the correct prediction ratio, without taking into account what is positive and negative.	100
Sensitivity	Sensitivity determines the ability of the system to correctly predict the condition of cases that actually have it, is the proportion of true positives.	100
Specificity	Specificity determines the system's ability to correctly predict the absence of the condition for cases that actually have it, is the proportion of true negatives.	100
Efficiency	Efficiency is the arithmetic mean of sensitivity and specificity, which suggests that if very sensitive to positives, it tends to generate many false positives and vice versa. The perfect decision method occurs when 100% sensitivity and 100% specificity	100
Positive Predictive Value	The positive predictive value is the proportion of true positives in relation to all positive predictions	0
Negative Predictive Value	The negative predictive value is the ratio of true negatives to all negative predictions	0

VII. Discussion of Results

This work shows several phases of technical developments, the final result of which demonstrates that the classification and grouping of signals occur, and also the determination of the variable (weights) for better adjustment. This was possible by applying the AR-ANN technique to the Wiener-data.

The processing of the signal by applying the Wiener filter, promotes the elimination of unwanted signal noise, with this, the formation of the data-Wiener database is of important application. Figure 16 already demonstrates the result of linear regression applied to the 4Hz database. In this case, the difference between the samples is noticeable, as there is no tendency of proximity of the bad signal with the best fit line.

Signal processing using ANN perceptron enables the grouping of different signals, thus having a methodology of identification by statistical parameters of fundamental importance. In this case, parameters such as standard deviation and variance were applied so that the signals were validated within a given analysis group. Figure 17 shows the cross-reference test of the samples, made randomly. This process makes it possible to obtain cross-results of the different samples between the reference and bad signals. Figure 18 shows the linear regression graph, whose signals processed by ANN, allow a view of the signals processed with the best condition of the variables (weight). Finally, figure 19 shows the result of network training, considering the application of the ANN technique with 164 epocs.

Signal processing using AR was applied in order to determine and analyze the behavior of the group analyzed by ANN, whose metrics make it possible to understand that the classification of the groups was successful and of very positive conditions. It is possible to take as an example the result of efficiency, the result of which demonstrates that a perfect decision method occurred, because the result was 100% of sensitivity and 100% specificity. Note also that the values of the parameters of positive predictive value and negative predictive value were equivalent to zero (0). This demonstrates that in this work, these types of signals were not found in the data-Wiener. This information can be seen in table 01.

Figure 20 shows the flow of database development to the definition of metrics.

Note that the information and its respective knowledge, generated by the processing of the Data-Wiener by AR-ANN, allows for more accurate decision-making. Figure 19 shows the potential of obtaining information and acquiring knowledge.

VIII. Conclusion

This work demonstrates that the combination of three mathematical concepts from different areas, Wiener Filter, ANN and AR, could be applied jointly in the fault detection analysis of a system. However, it also demonstrates a result of variables (weights) of better fit for the network. This multidisciplinary ideological concept enables the application in the analysis of rapid needs and online systems (real time), as it contributes with comparison parameters (weights) for standardization of future data. This type of multidisciplinary technique enables several applications in different areas of engineering, such as aerospace, mechanics, civil, among others.

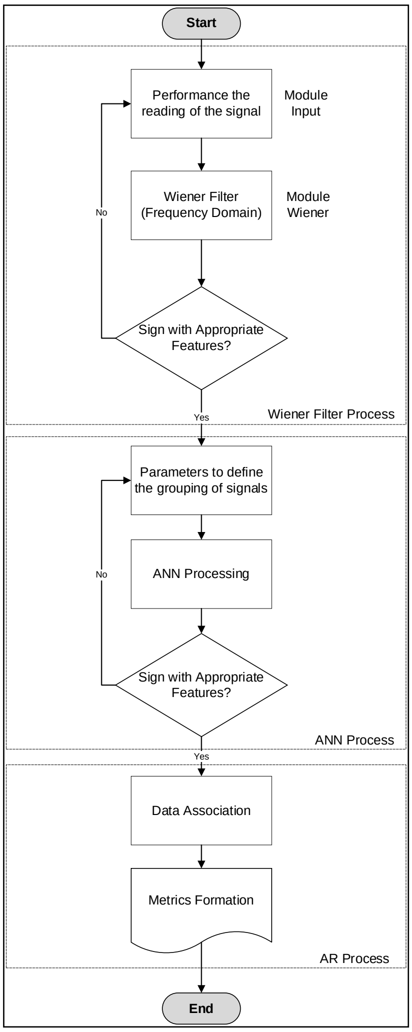


Figure 20. Multidisciplinary Development Flow

Concepts of this type can also be applied to predictive maintenance techniques, considering the principle of embedded systems and cybersystems, meeting the needs of industry 4.0.

Considering that the objective of this work was to develop a multidisciplinary method that could determine the failures of a coupled gear system, it is possible to determine that the result was obtained successfully, efficiently and accurately.

Acknowledgment

The authors thank the Paula Souza Center - CPS for their support and encouragement of research in FATEC's. They also thank the Laboratory of Complex Systems (Sisplexos) of the Institute of Chemistry - UNESP, for providing the physical space, the computer park and the financial support of the São Paulo State Research Support Foundation (FAPESP), through the process 2023/00861-8 and to the National Council for Scientific and Technological Development (CNPq), through process 301401/2022-5.

References

[1] Chen, Y., Li, Y., Computational Intelligence Assisted Design: In *Industrial Revolution 4.0*, CRC Press, Boca Raton, 2018. 527p.

[2] Wojciechowska, W. S., “Technical System Maintenance: Delay-Time-Based Modelling”, *Springer Nature*, Gewerbestrasse, 2019. 361p.

[3] Anand, A., Ram, M., “System Reliability Management: Solutions and Technologies”, *CRC Press, Boca Raton*, 2019. 277p.

[4] Silbestri, G., Verona, F. B., Innocenti, M., Napolitano, M., “Fault detection using neural networks”, *Proceedings of 1994 IEEE International Conference on Neural Networks* (ICNN'94), v.6, n.1, 3796-3799p., 1994. doi: 10.1109/ICNN.1994.374815.

[5] Heo, S., Lee, J.H., “Fault detection and classification using artificial neural networks”, *IFAC-PapersOnLine*, v. 51, n.18, 470-475p., 2018. doi.org/10.1016/j.ifacol.2018.09.380.

[6] Oliveira, D.C.; Chavarette, F.R.; Lopes, M.L.M., “Damage diagnosis in an isotropic structure using an artificial immune system algorithm”. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, v. 41, p. 485-492, 2019. doi:10.1007/s40430-019-1971-9.

[7] Outa, R.; Chavarette, F.R.; Mishra, V.N.; Gonçalves, A.C.; Roefero, L.G.P.; Moro, T.C., “Prognosis and fail detection in a dynamic rotor using artificial immunological system”. *Engineering Computations*, v. 37, n.9, p. 3127-3145, 2020. doi.org/10.1108/EC-08-2019-0351.

[8] Collins, J.A., Busby, H.R., Staab, G.H., “Mechanical Design Of Machine Elements And Machines: A Failure Prevention Perspective”, ed.2, *John Wiley & Sons*, Hoboken, 2010. 912p.

[9] Vullo, V., “Gears: Geometric and Kinematic Design”, *Springer Nature*, Switzerland, 2020. 880p.

[10] Smith, J.D., “Gear Noise and Vibration”, ed.2, *Marcel Dekker*, New York, 2003. 318p.

[11] Palmer, D., Fish, M., “Evaluation of Methods or Calculating Effects of Tip Relief on Transmission Error, Noise and Stress in Loaded Spur Gears”, *Gear Technology*, 01-12p. 2012.

[12] Park. S., Kim, S., Choi, J., “Gear Fault Diagnosis Using Transmission Error And Ensemble Empirical Mode Decomposition”, *Mechanical Systems and Signal Processing*, v.108, n.1, 262–275p. 2018. DOI.org/10.1016/j.ymssp.2018.02.028.

[13] De Silva, C.W., “Vibration: Fundamentals and Practice”, ed.3, *CRC Press*, Boca Raton, 2006. 1061p.

[14] Meirovitch, L., “Principles and Techniques of Vibrations”, *Prentice Hall*, Inc., Upper Saddle River, New Jersey 07458, 1997. 694.p.

[15] Inman, D.J., “Engineering Vibration”, ed.3, *Pearson Education*, Upper Saddle River, 2014. 720p.

[16] Crocker, M. J., “Handbook of Noise and Vibration Control”, *Hoboken, John Wiley & Sons*, 2007. 1569p.

[17] Ingle, V.K., Proakis, J.G., “Digital Signal Processing Using MATLAB”, 3 ed., *Cengage Learning*, Stamford, 2012. 671p.

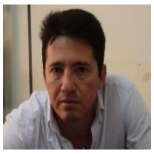
[18] Oppenhem, A.V., Schafer, R. W., Buck, J.R., “Discrete-Time Signal Processing”, 2ed., *Prentice Hall*, New Jersey, 1999. 897p.
[19] Hussain, Z.M., Sadik, A.Z., O´Shea, P., “Digital Signal Processing: An Introduction with MATLAB and Applications”, *Springer-Verlag, Heidelberg*, 2011.
[20] Lyons, R.G., “Streamlining Digital Signal Processing: A Tricks of the Trade Guidebook”, *John Wiley & Sons*, New Jersey, 2012. 466p.
[21] Outa, R.; Chavarette, F.R.; Toro, P.F.; Gonçalves, A.C.; Santos, L.H., “Prognosis and Detection of Experimental Failures in Open Field Diesel Engines Applying Wieners Artificial Immunological Systems”. *Journal of Applied and Computational Mechanics*, v. 7, p. 1-12, 2021. doi: 10.22055/JACM.2020.34972.2525.
[22] Holman, J.P., “Experimental Methods for Engineers”, 8ed., *McGraw-Hill Companies*, New York, 2012. 769p.
[23] Ye, N., !The Handbook of Data Mining”, *Lawrence Erlbaum Associates*, Mahwah, 2003. 722p.
[24] Wang, L., Fu, X., “Data Mining with Computational Intelligence”, *Springer-Verlag*, Heidelberg, 2005. 280p.
[25] Han, J., Kamber, M., Pei, J., “Data Mining Concepts and Techniques”, *Elsevier*, Waltham, 2012. 740p.
[26] Bigus, J.P., “Data Mining with Neural Networks: Solving Business Problems from Application Development to Decision Support”, *MacGraw-Hill Companies*, New York, 1996., 237p.
[27] Felici, G., Vercellis, C., “Mathematical Methods for Knowledge Discovery and Data Mining”, *IGI Global*, Convent Garden, 2008. 394p.
[28] Da Silva, I.N., Spatti, D.H., Flauzino, R.A., Liboni, L.H.B., Alves, S.F.R., “Artificial Neural Networks: A Pratical Course”, *Springer International*, Switzerland, 2017. 309p.
[29] Srisook, N., Tuntoolavest, O., Danphitsanuparn, P., Pattana-anake, V., Joseph, F.J.J. “Convolutional Neural Network Based Nutrient Deficiency Classification in Leaves of Elaeis guineensis Jacq”, *International Journal of Computer Information Systems and Industrial Management Applications*, v. 14, p. 019-027. 2022.

Author Biographies

Paulo Roberto Tebon is currently a student at the Faculty of Technology of Araçatuba, Araçatuba, São Paulo, Brasil.

Roberto Outa Graduated as a mechanical technologist in production processes (1992) from the College of Technology of São Paulo, specialized in Project Management (2011) from the University Toledo, Master in Mechanical Engineering (2014) from UNESP campus Solteira - FEIS, research project "Study and analysis of the sound absorption coefficient of sugarcane fiber composite", PhD in Mechanical Engineering (starting in 2014 - in activity) UNESP campus Solteira - FEIS, research project "Chaotic Motion suppression of a dynamic rotor using the Linear Optimal control". Since 2010 is a professor in industrial maintenance for Fatec Araçatuba / SP in the course of biofuels. Acting as project manager and product development engineer in private industry by developing and implementing concepts of engineering in production and products.

Fábio Roberto Chavarette is graduated in Bachelor of Computer Science from School of Engineering of Piracicaba (1996), master's degree in Applied Physics from Sao Paulo State University (2002), doctor in Mechanical Engineering from Universidade Estadual de Campinas (2005), Post-Doctorate (2010) in the Department of Statistics, Applied Mathematics and Computing, from Sao Paulo State University. Currently, is an Associate Professor in the Department of Engineering, Physics and Mathematics (DEFM) of the Institute of Chemistry (IQ), of the Universidade Estadual Paulista "Julio de Mesquita Filho"



Aparecido Carlos Gonçalves holds a degree in Mechanical Engineering from the University of São Paulo (1989), a master's degree in Metallurgical Engineering [S. Carlos] from the University of São Paulo (1993) and a PhD in Mechanical Engineering from the University of São Paulo (1997). He is currently a FULL Professor at University of State of São Paulo ? UNESP, advisor to the journals: Journal of Materials Science, Journal of Thermoplastic Composite Materials, Industrial Lubrication and Tribology. He has experience in Mechanical Engineering, with emphasis on Machines, Engines and Equipment, working mainly on the following topics: oil analysis, vibration analysis, predictive maintenance, particle analysis and mechanical systems projects. Coordinated research projects at FINEP, CNPq (Universal Range C), FAPESP (5), Fundunep (Fundation for the Development of Unesp), Fepisa (Teaching, Research and Extension Foundation of Ilha Solteira) and Proex (Pro-rector of Extension University of Unesp). Doctoral supervisor.

Sandro da Silva Pinto graduation at Production Engineering from Universidade Federal de São Carlos (1997) and master's at Production Engineering from Universidade Federal de São Carlos (2000). He is currently professor horista graduação at Centro Universitário de Lins, coordenador pós-graduação lato sensu at Centro Universitário de Lins, professor tutor at Centro Universitário de Lins, professor associado ii at Faculdade de Tecnologia de Lins and coordenador tecnologia em gestão da qualidade at Faculdade de Tecnologia de Lins. Has experience in Production Engineering, focusing on Production Engineering, acting on the following subjects: indústria sucroalcooleira, gestão, logística, qualidade; padronização and gestão por processos.

Samuel Stabile graduate at Tecnologia em Processamento de Dados from Faculdade de Tecnologia de Birigui (1996) and master's at Engenharia de Produção - São Carlos from Universidade de São Paulo (2001). Has experience in Production Engineering, acting on the following subjects: sistemas de informação, desconexão entre usuários e desenvolvedores, competitividade, gerência and controle.