electronics

Article

Advanced Fault-Detection Technique for DC-Link Aluminum Electrolytic Capacitors Based on a Random Forest Classifier

Acácio M. R. Amaral [1,2,*], Khaled Laadjal [2] and Antonio J. Marques Cardoso [2]

1 Polytechnic Institute of Coimbra, Coimbra Institute of Engineering, Rua Pedro Nunes—Quinta da Nora, P-3030-199 Coimbra, Portugal
2 CISE-Electromechatronic Systems Research Centre, University of Beira Interior, Calçada Fonte do Lameiro, P-6201-001 Covilhã, Portugal
* Correspondence: amramaral@ieee.org or acacio@isec.pt

Abstract: In recent years, significant technological advances have emerged in renewable power generation systems (RPGS), making them more economical and competitive. On the other hand, for the RPGS to achieve the highest level of performance possible, it is important to ensure the healthy operation of their main building blocks. Power electronic converters (PEC), which are one of the main building blocks of RPGS, have some vulnerable components, such as capacitors, which are responsible for more than a quarter of the failures in these converters. Therefore, it is of paramount importance that the design of fault diagnosis techniques (FDT) assess the capacitor's state of health so that it is possible to implement predictive and preventive maintenance plans in order to reduce unexpected stoppage of these systems. One of the most commonly used capacitors in power converters is the aluminum electrolytic capacitor (AEC) whose aging manifests itself through an increase in its equivalent series resistance (ESR). Several advanced intelligent techniques have been proposed for assessing AEC health status, many of which require the use of a current sensor in the capacitor branch. However, the introduction of a current sensor in the capacitor branch imposes practical restrictions; in addition, it introduces unwanted resistive and inductive effects. This paper presents an FDT based on the random forest classifier (RFC), which triggers an alert mechanism when the DC-link AEC reaches its ESR threshold value. The great advantage of the proposed solution is that it is non-invasive; therefore, it is not necessary to introduce any sensor inside the converter. The validation of the proposed FDT will be carried out using several computer simulations carried out in Matlab/Simulink.

Keywords: fault diagnosis technique; aluminum electrolytic capacitors; short-time least square Prony's (STLSP); random forest classification

Citation: Amaral, A.M.R.; Laadjal, K.; Cardoso, A.J.M. Advanced Fault-Detection Technique for DC-Link Aluminum Electrolytic Capacitors Based on a Random Forest Classifier. *Electronics* **2023**, *12*, 2572. https://doi.org/10.3390/electronics12122572

Academic Editors: Ryad Zemouri, Mélanie Lévesque and Arezki Merkhouf

Received: 12 May 2023
Revised: 27 May 2023
Accepted: 6 June 2023
Published: 7 June 2023

1. Introduction

The monitoring of power electronic (PE) systems, converters, and components is crucial in applications that require reliability and safety, including electric aircraft, electric automobiles, wind turbines, etc. Condition monitoring (CM) enables the prediction of potential failures and events, allowing for the implementation of appropriate preventive and predictive maintenance approaches to keep these systems functioning. In [1], the authors provide a comprehensive review of semiconductor device condition monitoring, while capacitors represent another class of reliability-critical components. Capacitors CM has been taken into account for various dc-link applications, including power factor correction (PFC) converters [2], adjustable speed drive (ASD) systems [3], and photovoltaic (PV) grid-connected inverters [4]. There are numerous implementation strategies, such as using voltage and current data from controllers or sampling signals from additional hardware circuits and sensors. Over the past two decades, several efforts [3–8] have been made to implement CM techniques for capacitors in dc-link applications, some requiring the removal of capacitors from the converters [9]. Others, such as real-online [8,10] and quasi-online systems, are implemented locally in a real system. Figure 1 shows the classification of the various CM techniques.

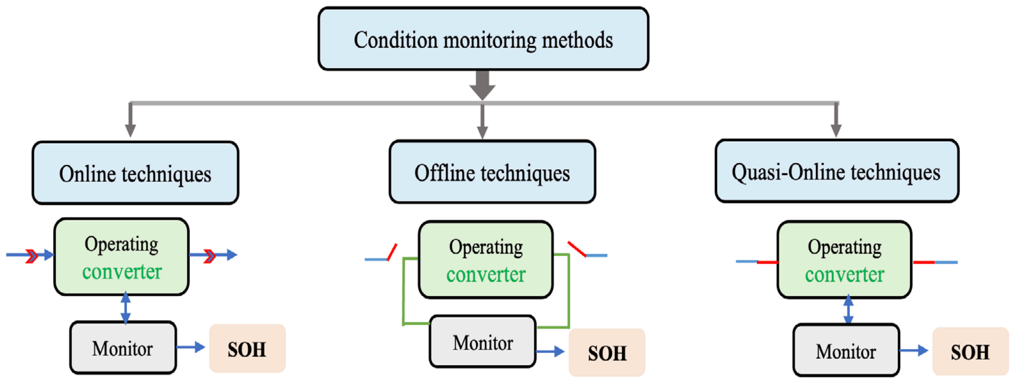

Figure 1. Classification of condition monitoring techniques.

The aluminum electrolytic capacitor (AEC) is commonly used in power systems due to its high volumetric efficiency and low cost. It is available in various sizes and capacities [11]. However, AECs must be connected in series and require safety resistors to maintain the necessary high voltage in power networks, leading to a more complex system that may result in additional damage when they fail [12,13]. AEC failure can be caused by several thermal and electrical issues, leading to the vaporization of its electrolyte and degradation of the oxide layer. This results in a significant decrease in capacitance (C) and an increase in equivalent series resistance (ESR) rates [6,13–16].

However, the effective lifespan of an AEC is shorter than that of converters, and it needs to be replaced once or twice during a converter's lifespan [14,15]. To monitor the degraded condition of the capacitor, condition monitoring methods focus on estimated values of capacitance (C) or equivalent series resistance (ESR). According to capacitor datasheets, an AEC's lifespan is over if the C value drops by 20% or the ESR doubles from its original rate [17]. Several investigative techniques have been proposed to examine the condition of AECs used in converters [18]. To avoid using a current sensor in the DC link, several techniques have been suggested [19]. In these methods, the capacitor current is obtained based on the circuit relationship to monitor the health of the AEC. However, all these techniques require additional hardware or complex algorithms.

Experimental techniques have been utilized to evaluate the condition of aluminum electrolytic capacitors (AECs) in converters [13–16,18,20,21]. Some approaches require the presence of a current sensor connected in series with the capacitor [14,15,18,20], which could increase the stray inductance and alter the capacitor voltage, as previously reported [20]. However, there are alternative methods that can replace the current sensor in the DC link [13,16,19,21]. These methods rely on the circuit relationship and the capacitor current to assess the AEC's condition. However, all of these techniques require either additional hardware or complex algorithms.

A recent study investigated the feasibility of using the dissipation factor (DF) as an indicator of the lifespan of AECs [22]. DF is commonly used to establish a threshold for determining whether electrolytic capacitors are still operational. It is the tangent of the angle formed between the impedance vector of the capacitor and the negative reactive axis, when the electrical parameters are represented as vectors in a complex plane. This angle, also known as the loss angle or impedance angle, is the counterpart of the angle between the voltage and current vectors of the capacitor. Therefore, by measuring the impedance angle, one can calculate the DF and hence the health status of the capacitor. In [23], the capacitance of a capacitor was determined by monitoring its discharge time in a converter supplying a motor load. On the other hand, in [24], the authors proposed a technique to calculate the capacitance of capacitors in a multimodule converter (MMC) by monitoring their discharge time in the bleeding resistors.

Traditional estimation methods demand complex systems, extra hardware, and strict requirements such as signal injections and sampling time. These issues can be addressed and a better performance can be achieved by employing advanced machine learning (ML) techniques. In conventional methods, the capacitor current is often obtained directly or indirectly through circuit relationships and switching states using a current sensor, which makes the estimation process more complicated and expensive. With the advancement of information technology, artificial intelligence (AI) methods have become popular and offer potential strategies for fault diagnosis, such as arc fault detection [25–29]. Innovative algorithms such as artificial neural networks and adaptive neuro-fuzzy inference systems (ANFIS) have been used to evaluate the health status of AECs [30]. ANFIS assesses the aging status of AECs in the converter by fitting curves to the valued parameters and actual capacitor factors. These techniques can monitor capacitor condition by using feedback data generated during usual and aging error conditions of the capacitor. However, research on AI algorithms based on capacitor estimation is still in its early stages, and a comprehensive estimation process has yet to be established.

Non-invasive FDT (NIFDT) design has recently become a hot topic due to the practical constraints associated with the introduction of sensors inside the converters. Some authors have proposed solutions to overcome these restrictions, which require additional hardware or complex algorithms. This paper aims to contribute with an NIFDT that, in addition to not being complex, does not require the use of additional hardware. The proposed solution combines a machine learning (ML) algorithm, the random forest classifier (RFC), with the short-time least squares Prony (STLSP) technique to assess Al-Cap state of health.

2. Short-Time Least Square Prony's Technique

The short-time least square Prony's (STLSP) technique is a high-resolution method used to perform an online estimation of the ESR parameter, allowing thus, a continuous evaluation of the AEC condition. The advantage of using the STLSP technique is that it can correctly determine and monitor all harmonic properties (frequency, amplitude, phase, and damping factor) from just a short record of a signal. By considering the signal $y(t)$ and its N complex samples, Prony's method estimates the sampled data by employing a linear combination of P complex exponential functions [8,10]:

$$\hat{y}(n) = \sum_{k=1}^{P} w_k m_k^{n-1} \tag{1}$$

with $w_k = A_k e^{j\varphi_k}$ and $m_k = e^{(\alpha_k + j2\pi f_k)T_s}$, T_s is the sampling time. The model parameters A_k, f_k, φ_k, and α_k represent, respectively, the unknown amplitude, frequency, phase angle, and damping factor of the kth component.

The aforementioned equation presents a challenging non-linear problem that can be effectively addressed through the application of Prony's method. In essence, Prony's method transforms the problem of parameter estimation from non-linear to linear, achieved by solving a linear system and calculating the roots of the polynomial. Consequently, this method establishes a homogenous linear difference equation with constant coefficients, where $a_0 = 1$.

$$\sum_{k=0}^{P} a_k y[n-k] = 0 \tag{2}$$

In the conventional Prony's method, it is assumed that the number of available data samples is equal to the unknown model parameters. Consequently, the linear difference Equation (2) can be represented in matrix form as follows:

$$Y.A = -\lambda \tag{3}$$

where $Y = \begin{bmatrix} y[P] & \cdots & y[1] \\ \vdots & \ddots & \vdots \\ y[2P-1] & \cdots & y[P] \end{bmatrix}$, $A = \begin{bmatrix} a_1 \\ \vdots \\ a_P \end{bmatrix}$, $\lambda = \begin{bmatrix} y[P+1] \\ \vdots \\ y[P] \end{bmatrix}$

The linear prediction parameters, a_k, which best fit the observed data, are determined by solving Equation (3). Subsequently, the linear prediction parameters are utilized to create a characteristic polynomial with roots, m_k, in the following manner:

$$f(m) = \sum_{k=0}^{P} a_k m^{P-k} \tag{4}$$

Consequently, the damping factor and frequency can be obtained directly from the roots, m_k, of Equation (1):

$$\alpha_k = \frac{\ln|m_k|}{T_s} \text{ and } f_k = \frac{1}{2\pi T_s} \tan^{-1}\left[\frac{\mathrm{Im}(m_k)}{\mathrm{Re}(m_k)}\right]$$

Finally, the roots m_k are utilized to write the P equations of (1) in a matrix form as:

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ m_1 & m_2 & \cdots & m_P \\ \vdots & \vdots & \vdots & \vdots \\ m_1^{P-1} & m_2^{P-1} & \cdots & m_P^{P-1} \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_P \end{bmatrix} = \begin{bmatrix} y(1) \\ \vdots \\ y(P) \end{bmatrix} \tag{5}$$

The complex parameters mk can be determined by solving (5), and consequently, the exponential amplitudes A_k and phase angles φ_k can be obtained using the following relationships:

$$A_k = |w_k| \text{ and } \varphi_k = \tan^{-1}\left[\frac{\mathrm{Im}(w_k)}{\mathrm{Re}(w_k)}\right]$$

On the other hand, in practice, the number of available data samples overrides the number of unknown parameters ($N > 2P$). In the over-determined data case, the linear difference should be as follows [8,10]:

$$\sum_{k=0}^{P} a_k y[n-k] = \varepsilon[n] \tag{6}$$

The available N data samples are used to rewrite (6) in a matrix form:

$$\begin{bmatrix} y[P] & \cdots & y[1] \\ \vdots & \ddots & \vdots \\ y[N-1] & \cdots & y[N-P] \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_P \end{bmatrix} = -\begin{bmatrix} y[P+1] \\ \vdots \\ y(N) \end{bmatrix} \tag{7}$$

The vector of the unknown parameters a_k is picked to minimize the linear prediction total squared error. The minimization can be solved by using the least square method. In addition, the estimation of the complex parameters w_k is turned also into a linear least square procedure.

$$\mathrm{M.W} = C \tag{8}$$

with:

$$M = \begin{bmatrix} 1 & \cdots & 1 \\ m_1 & \cdots & m_P \\ \vdots & \cdots & \vdots \\ m_1^{N-1} & \cdots & m_P^{N-1} \end{bmatrix}, \quad W = \begin{bmatrix} w_1 \\ \vdots \\ w_P \end{bmatrix}, \quad C = \begin{bmatrix} y(1) \\ \vdots \\ y(N) \end{bmatrix} \tag{9}$$

It is crucial to keep in mind that the objective is to estimate and monitor the parameters C (capacitance) and ESR (equivalent series resistance) since they offer valuable insights into the potential failure of the capacitor. Any variations in the C and ESR parameters are manifested in the ratio between the capacitor voltage ripple and the current ripple. This ratio, at any given moment, equals the impedance of the capacitor. In the low-frequency range, the impedance is primarily governed by C, while in the high-frequency range, it is dominated by ESR. Therefore, the following formula may be used to calculate ESR:

$$ESR = \frac{V_{fsw}}{I_{fsw}} \tag{10}$$

where V_{fsw} and I_{fsw} are the amplitudes of the switching frequency harmonics. All these spectral components are permanently present in the capacitor voltage ripple and current ripple.

3. Non-Invasive Fault-Detection Technique

Climate change has forced many countries to accelerate the transition to a sustainable energy matrix, therefore, with a greater predominance of renewable energy sources.

On the other hand, some fundamental elements of renewable power generation systems (RPGS) are the power converters [31,32], which must present very high reliability. One of the most vulnerable components of power converters is the aluminum electrolytic capacitor (AEC), which is responsible for more than 25% of their failures [11,14]. Therefore, it is extremely important to design fault diagnosis techniques (FDT) that assess the AECs state of health during power converter operation, so that RPGS good operation can be guaranteed.

The aging of AECs produces an increase in their equivalent series resistance (ESR). According to the manufacturers [33], the ESR value can double in relation to its initial value when the capacitor reaches its end-life limit. Therefore, ESR can be used as a good failure indicator.

One of the most common methods for estimating the ESR value is by computing the ratio between the voltage and current amplitudes in the AECs at the converter operating frequency [9,34]. The previously described solution imposes the introduction of a current sensor inside the converter, which in many applications is not feasible. Furthermore, the current sensor introduces a resistive and inductive effect in the AEC branch, which is undesirable in applications that operate at high frequency and, at the same time, may require the converter redesign [35].

This article presents a solution that does not require the use of sensors inside the converter, as the proposed FDT does not involve current measurement in the AEC branch. The non-invasive FDT (NIFDT) uses an approach that combines the STLSP algorithm with the random forest classification (RFC), as can be seen in the general scheme shown in Figure 2.

The implementation of the NIFDT requires two phases, the first phase requires the training of the ML model (training stage), so that in a second phase it can operate in the final application.

The STLSP algorithm is used to process the capacitor voltage amplitude at the converter switching frequency (A_VC), which represents one fundamental attribute that is used in both the training and final application phases (Figure 2).

The ML algorithm, the RFC, will be trained to identify when the ESR value exceeds a predefined threshold. In this paper, the threshold corresponds to 1.5 times the ESR value of a sound capacitor. The training step is performed on a converter subject to a wide range of operating conditions, namely, different duty cycles (D_C) and load resistances (R_L). In this process, sound capacitors (ESR_{sound}) and age capacitors ($2 \times ESR_{sound}$) will be used. The previous procedure will create a dataset that will be fundamental for the RFC model to learn how to predict AEC health status (Figure 2).

After training, both the RFC model and the STLSP algorithm must be incorporated into the converter MCU (microcontroller unit). Finally, in the final application, during converter operation, the state of health of the AEC will be evaluated using the RFC model, which will have as inputs the value of A_Vc, calculated through the STLSP algorithm, and the values of R_L and D_C (Figure 2).

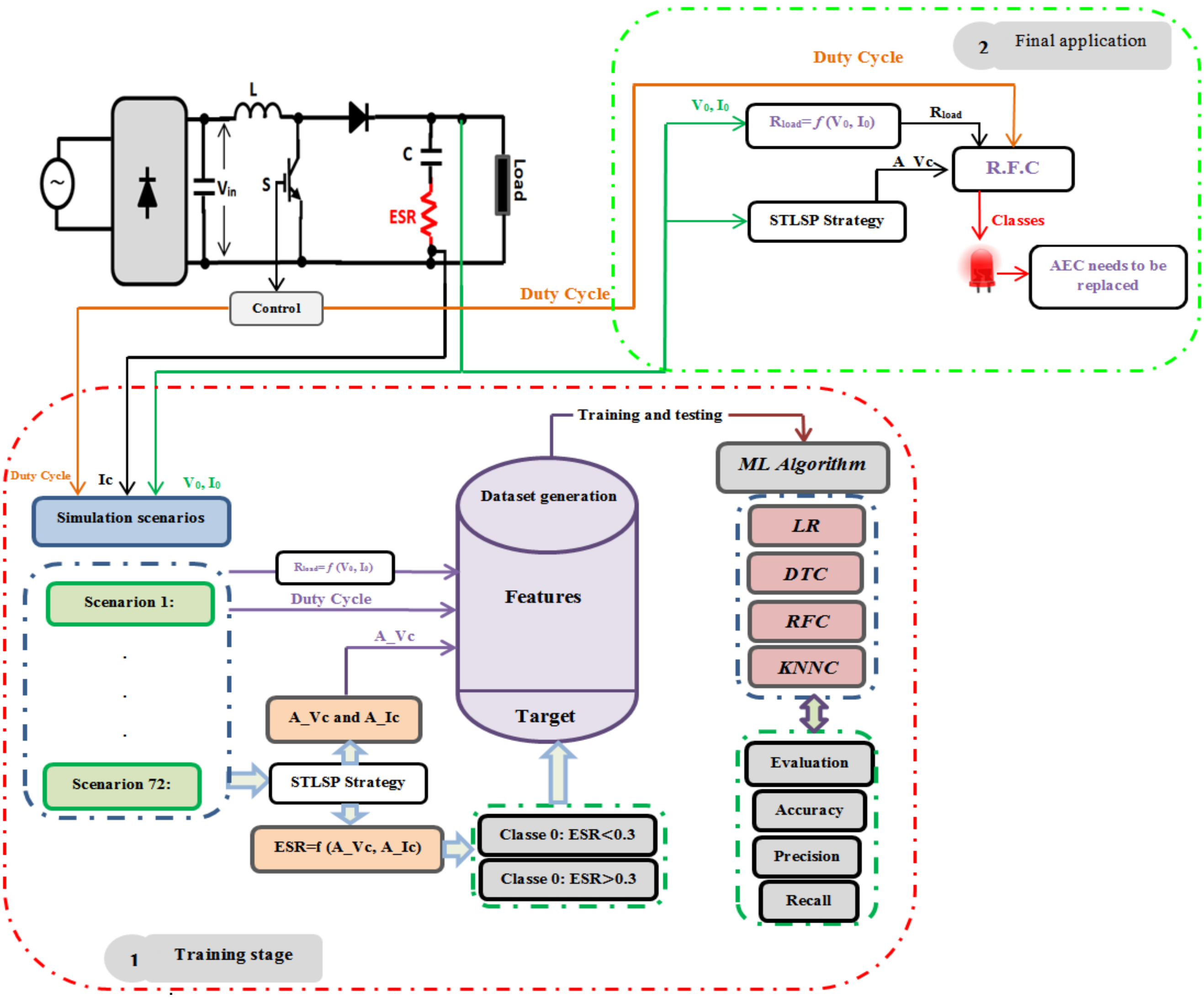

Figure 2. General scheme of the proposed NIFDT.

4. Simulation and Collecting the Data

The methodology used in this work follows the following steps:

1. Dataset creation;
2. Preprocessing;
3. Feature selection;
4. ML selection;
5. ML algorithm training, testing, and evaluating;
6. Application of NIFDT to completely new scenarios.

4.1. Dataset Creation

Before proceeding with ML model training, it is necessary to create two datasets. The first dataset will be used to select the best attributes, identify the best-performing ML algorithm, and train and test the ML model. The second dataset will be used to evaluate the performance of the ML model against a completely new dataset, in order to assess whether the ML model can be generalized.

To create both datasets, it was necessary to simulate the power converter, a boost converter (Figure 2), as well as to implement the STLSP algorithm in Matlab/Simulink simulation platform. Table 1 shows the simulated scenarios within the first dataset.

Table 1. Training and test scenarios.

Scenarios	R_{Load} (Ω)	Duty Cycle (%)	ESR (Ω)
1–9	25, 50, 75, 100	10, 20, 30, 40, 50, 60, 70, 80 and 90	0.2
10–18			0.4
19–27			0.2
28–36			0.4
37–45			0.2
46–54			0.4
55–63			0.2
64–72			0.4

After carrying out the simulations, the maximum possible attributes were extracted with the aid of the STLSP algorithm. In this procedure, it was taken into account that the ML model cannot impose the introduction of sensors inside the system. Thus, none of the attributes resulting from the capacitor current can be used. Consequently, the remaining attributes are:

1. A_VC—the amplitude of capacitor voltage (v_C) at the converter switching frequency (f_{sw}).
2. D_VC—the damping factor of v_C at f_{sw}.
3. P_VC—the phase angle of v_C at f_{sw}.
4. F_VC—the estimated f_{sw} of v_C.
5. R_L—the load resistance.
6. D_C—the duty cycle.

Figure 3 shows the simulation results (Table 1), with all attributes, as well as the ESR value, which was obtained through Equation (10) after extracting A_VC and capacitor current amplitude (A_IC) at f_{sw}. The A_VC and A_IC were obtained through STLP techniques making the training process automatic without any manual intervention.

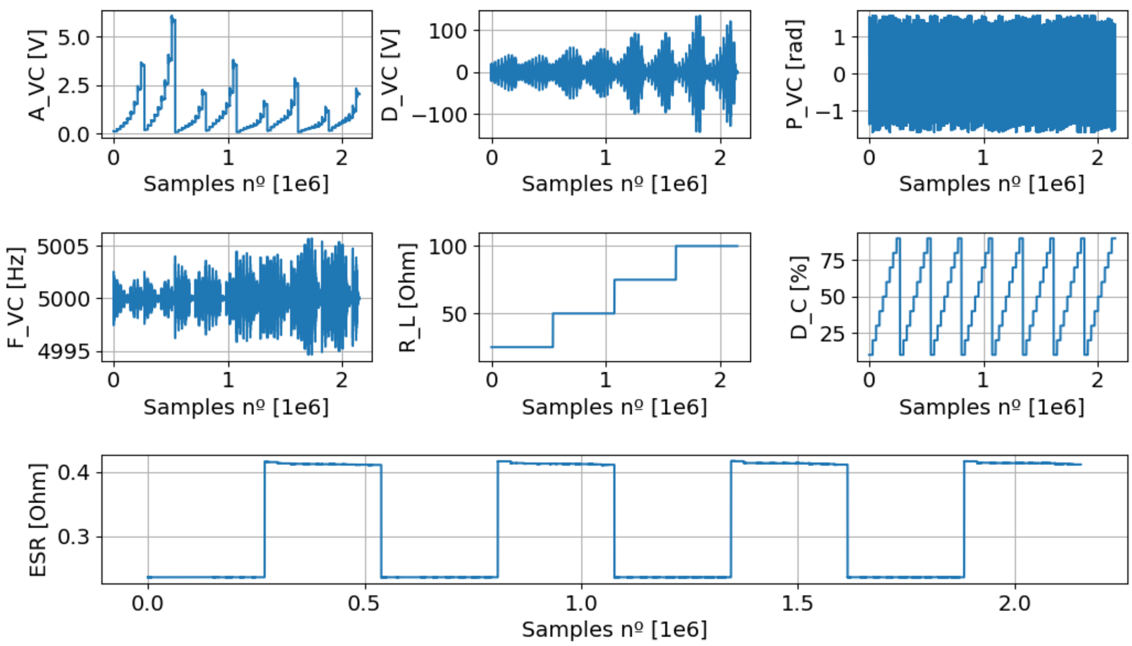

Figure 3. Initial dataset: all initial attributes and ESR estimated using the STLP technique.

4.2. Preprocessing

Before proceeding to the next phase, it was necessary to preprocess the dataset, namely, to reduce the number of samples, filter some attributes, and convert the ESR values into two classes.

Downsampling proves to be fundamental to reducing the computational cost, hence it is possible to significantly reduce the processing time of tasks 3–5. Filtering is equally fundamental for the A_VC, D_VC, P_VC, and F_VC attributes, as can be seen in Figure 3. As the ML algorithm is a classifier, it is mandatory to create two classes for training:

1. Class 0—ESR $< 0.3\ \Omega$ (sound capacitor);
2. Class 1—ESR $\geq 0.3\ \Omega$ (aged capacitor).

As a result of the described processing, the dataset represented in Figure 4 was generated.

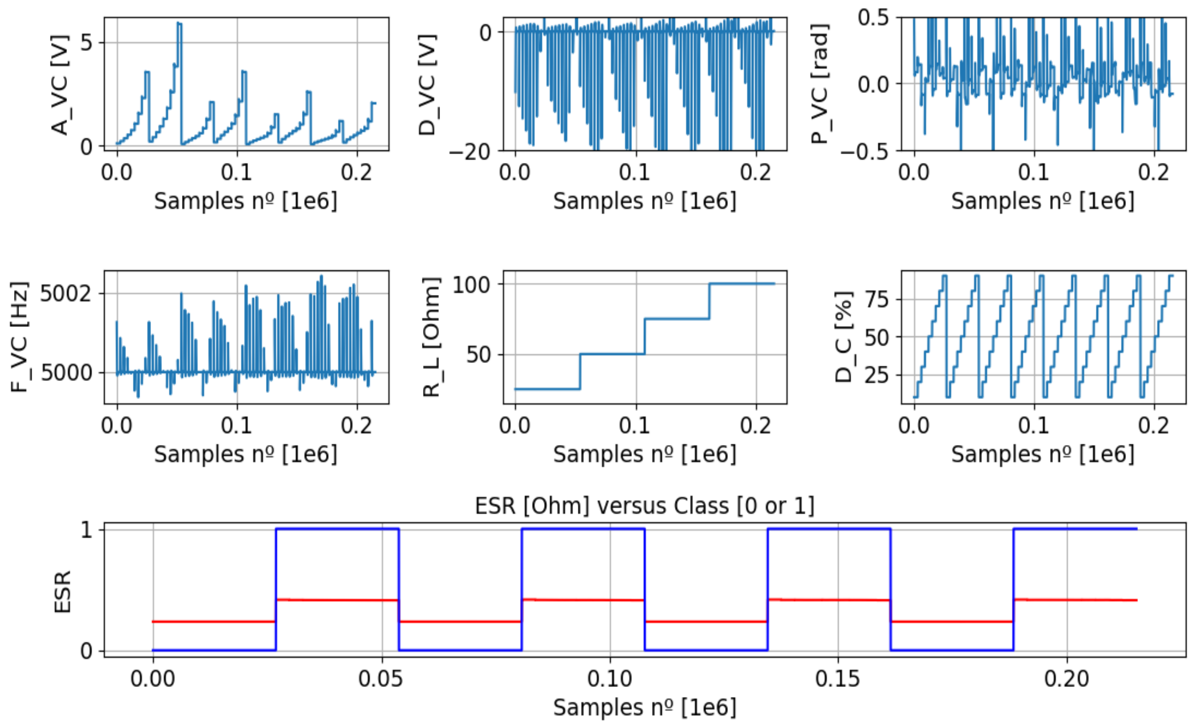

Figure 4. Post-processed dataset.

5. Feature Selection

The performance of the ML model does not depend solely on the ML algorithm but also on the attributes used to perform the predictions. It is therefore essential to choose the most appropriate attributes, that is, those that improve the model's performance. It should be noted that the selection of irrelevant attributes increases the model's complexity, increases the computation time, and can introduce noise, which can lead to overfitting and, therefore, reduce the model's performance [36–38].

Feature selection methods can be subdivided into unsupervised and supervised methods. The first ones do not need the target to select the attributes; the second subset needs the target.

In this paper, two unsupervised feature selection methods were used:

1. Pearson correlation threshold—eliminate attributes strongly related to each other (Figure 5), and a limit of 0.85 was imposed. The Pearson correlation (r) between two variables X and Y can be computed using (11):

$$r = \frac{n \times \sum(X \times Y) - (\sum X) \times (\sum Y)}{\sqrt{\left(n \times \sum X^2 - (\sum X)^2\right)} \times \sqrt{\left(n \times \sum Y^2 - (\sum Y)^2\right)}} \quad (11)$$

where n represents the number of samples.

2. Variance threshold—eliminates attributes that have low variance (Figure 6), that is, the ones that are approximately constant and very close to the mean. A threshold limit of 0.02 was considered. The variance (σ^2) of n samples of the quantity X, measures the variability from the mean ($\overline{X}$), and can be computed as follows:

$$s^2 = \frac{\sum(X - \overline{X})^2}{n} \quad (12)$$

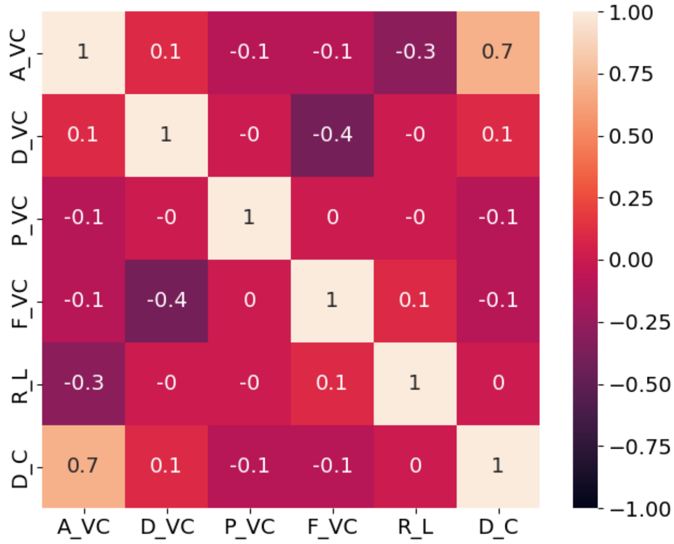

Figure 5. Attributes correlation matrix.

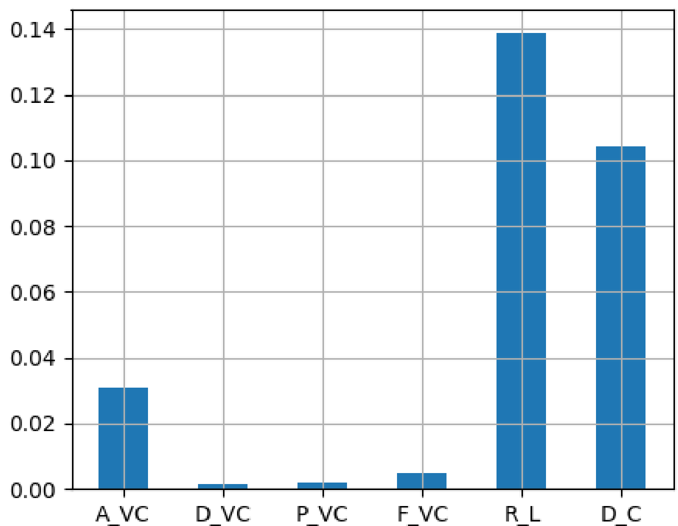

Figure 6. Attributes variance after scaling.

The previous figure shows that the correlation between all attributes is lower than the selected threshold, which means that this method does not eliminate any attribute.

Before computing the variance of each attribute, it was necessary to scale and transform each one to an interval between zero and one. Figure 6 shows the variance of each attribute.

Through the analysis of the previous figure, it is possible to select three attributes:

1. A_VC—the amplitude of capacitor voltage (v_C) at the converter switching frequency (f_{sw}).
2. R_L—the load resistance.
3. D_C—the duty cycle.

6. Machine Learning Algorithm Selection

In order to select the ML algorithm that best fits the problem under analysis, four of the most common classification algorithms were evaluated: logistic regression (LR), decision trees classification (DTC), random forest classification (RFC), and K-nearest neighbors classification (KNNC). For this purpose, three of the most commonly used metrics to evaluate the performance of ML classification algorithms were used: accuracy (13), precision (14), and recall (15).

To better understand the concepts of accuracy, precision, and recall, it is important, at first, to identify the samples related to sound capacitors (ESR < 0.3) as "negative samples" and the samples related to the aged capacitors (ESR > 0.3) as "positive samples".

Consequently, the models' outcome can be subdivided into:

1. True positive (TP)—represents the number of correct predictions that identify an aged capacitor.
2. False positive (FP)—represents the number of incorrect predictions that identify an aged capacitor.
3. True negative (TN)—represents the number of correct predictions that identify a sound capacitor.
4. False negative (FN)—represents the number of incorrect predictions that identify a sound capacitor.

In the cases of a TP or FP the ML model returns a '1', and in the case of a TN or FN the model will return a '0'.

Hence, it is possible to define accuracy as:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{13}$$

Accuracy is considered a good metric for balanced datasets, that is, for datasets where the number of positive samples is almost equal to the number of negative samples, as is the case with the Figure 4 dataset.

Precision represents how accurate the model is when it predicts a '1', and can be represented by the following equation:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{14}$$

Recall represents how accurate the model is when the true class is predicted as a '1', and can be represented by the following equation:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{15}$$

In order to assess the model's performance, 25 different datasets were created, containing random samples. For each of the 25 tests, the models were trained on just 1% of the data and tested on the remaining 99%.

Figure 7a shows the accuracy of the ML algorithms considering each individual dataset and Figure 7b shows the average value of the computed accuracies resulting from the 25 tests.

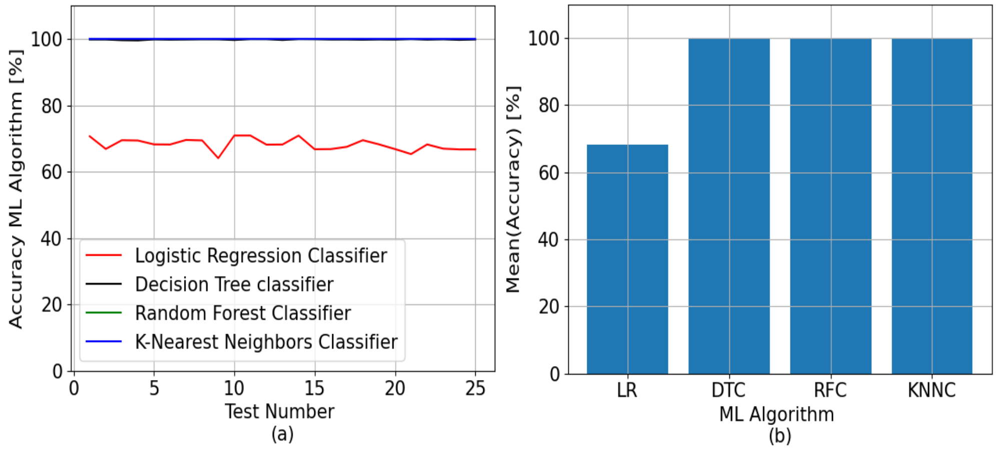

Figure 7. Accuracy of ML algorithms: (**a**) each individual dataset; (**b**) average value.

The same procedure was carried out regarding precision computation (Figure 8).

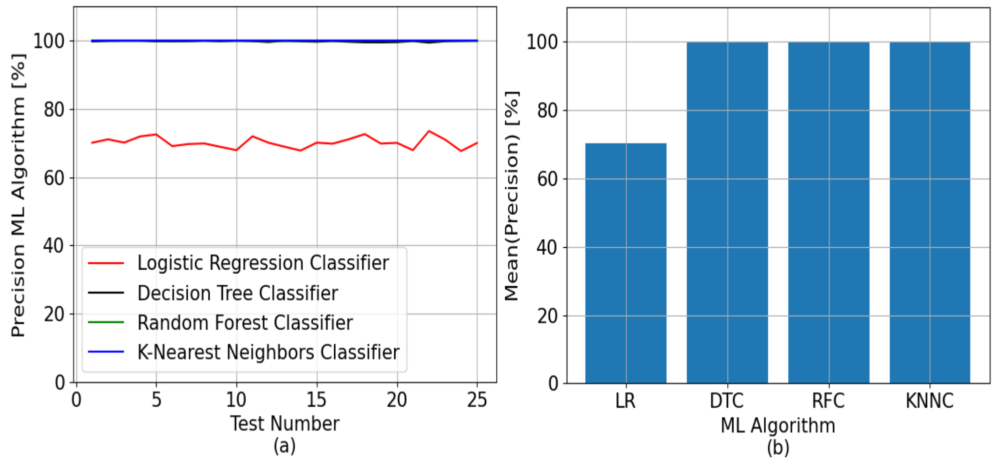

Figure 8. Precision of ML algorithms: (**a**) each individual dataset; (**b**) average value.

The same procedure was carried out regarding the recall computation (Figure 9).

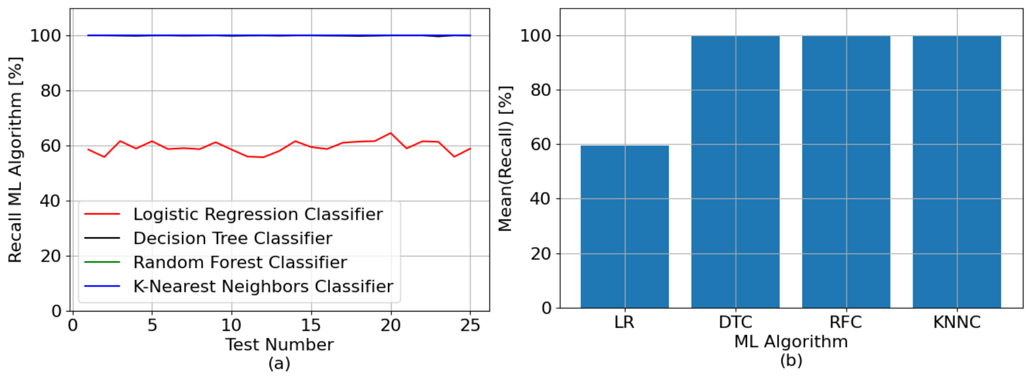

Figure 9. Recall of ML algorithms: (**a**) each individual dataset; (**b**) average value.

The previous figures show an overlap of the curves related to the DTC, RFC, and KNNC algorithms. However, when zoom is applied, it is possible to verify that the DTC algorithm performs slightly worse than RFC and KNNC algorithms.

The three metrics lead to the same conclusion, that is, the best ML models are the DTC, RFC, and KNNC. Although the DTC model requires less computational power, it has a greater tendency to overfit. Therefore, the DTC, the RFC, and the KNNC turn out to be the best ML algorithms.

7. Machine Learning Algorithm Training, Testing, and Evaluation

For ML algorithm training, the dataset represented in Figure 4 was used. At first, the training dataset was subdivided into a training set and a test set. The training set represents only 10% of the training dataset, with samples chosen randomly.

7.1. LR Evaluation

Initially, the training and testing stage of the LR model was performed, in order to confirm the conclusions of the previous section.

Figure 10 shows the predicted results of the LR model.

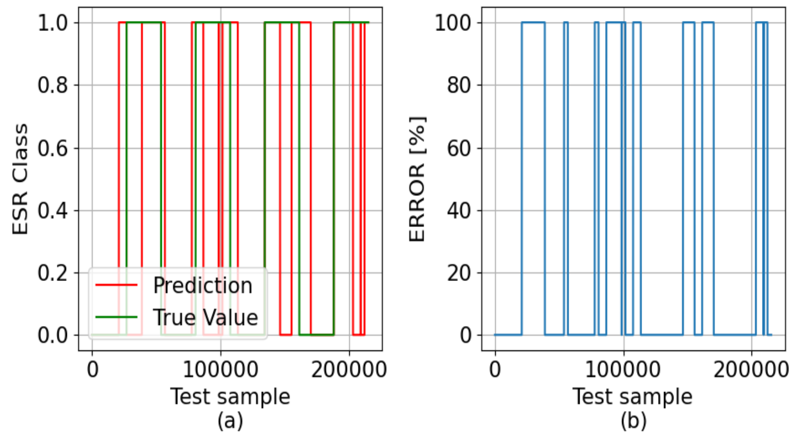

Figure 10. LR model predictions (dataset of Figure 4): (**a**) ESR Class; (**b**) Error.

The previous figure shows an accuracy of 67%, a precision of 70%, and a recall of 59%. These values are very similar to those obtained in the previous section. Thus, it can be concluded that the LR model is not suitable for the problem under analysis.

7.2. DTC Evaluation

Then, the training and testing stage of the DTC model was carried out, with the predictions shown in Figure 11.

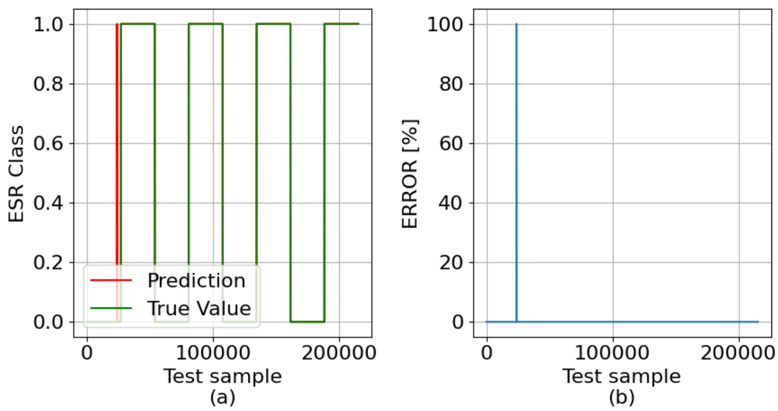

Figure 11. DTC model predictions (dataset of Figure 4): (**a**) ESR Class; (**b**) Error.

The previous figure clearly shows that the accuracy, precision, and recall values are close to 100%.

7.3. RFC Evaluation

Thereafter, RFC model training and testing were carried out. Figure 12 shows the prediction results.

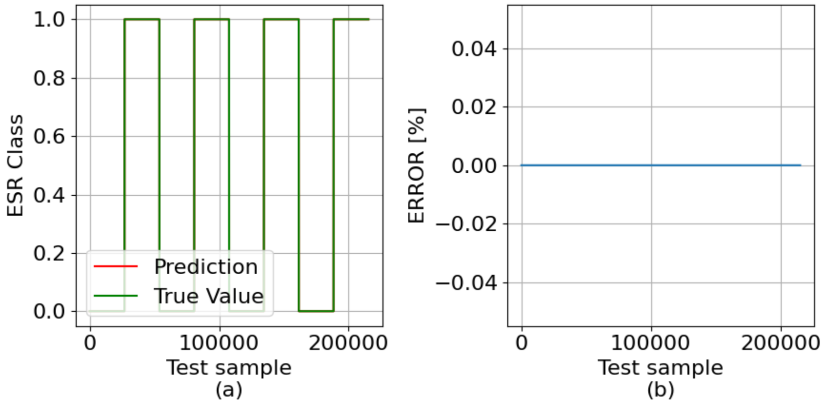

Figure 12. RFC model predictions (dataset of Figure 4): (**a**) ESR Class; (**b**) Error.

The previous figure clearly shows that the accuracy, precision, and recall values are 100%.

7.4. KNNC Evaluation

Finally, KNNC model training and testing were carried out. Figure 13 shows the prediction results.

The previous figure clearly shows that the accuracy, precision, and recall values are 100%. The results obtained in this section reiterate the conclusions of Sections 5 and 6: the best attributes are A_VC, R_L and D_C and the best ML algorithms are DTC, RFC, and KNNC.

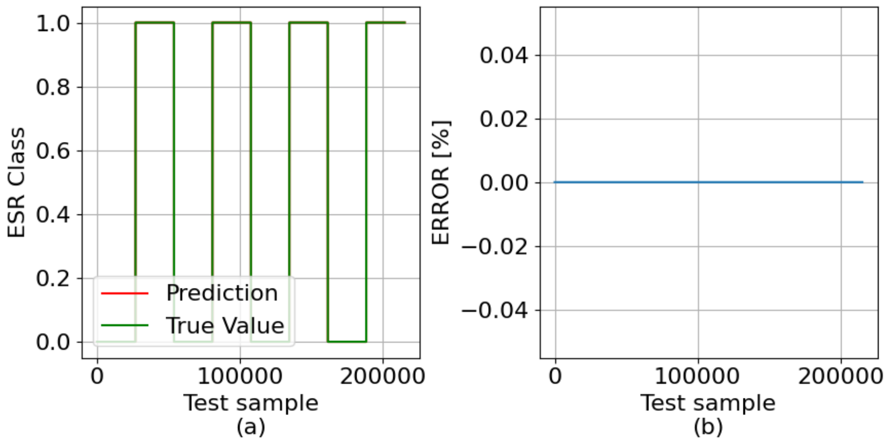

Figure 13. KNNC model predictions (dataset of Figure 4): (**a**) ESR Class; (**b**) Error.

8. Application of NIFDT to Completely New Scenarios

In order to evaluate the ability of the models to be generalized, a completely new dataset was created (Figure 14).

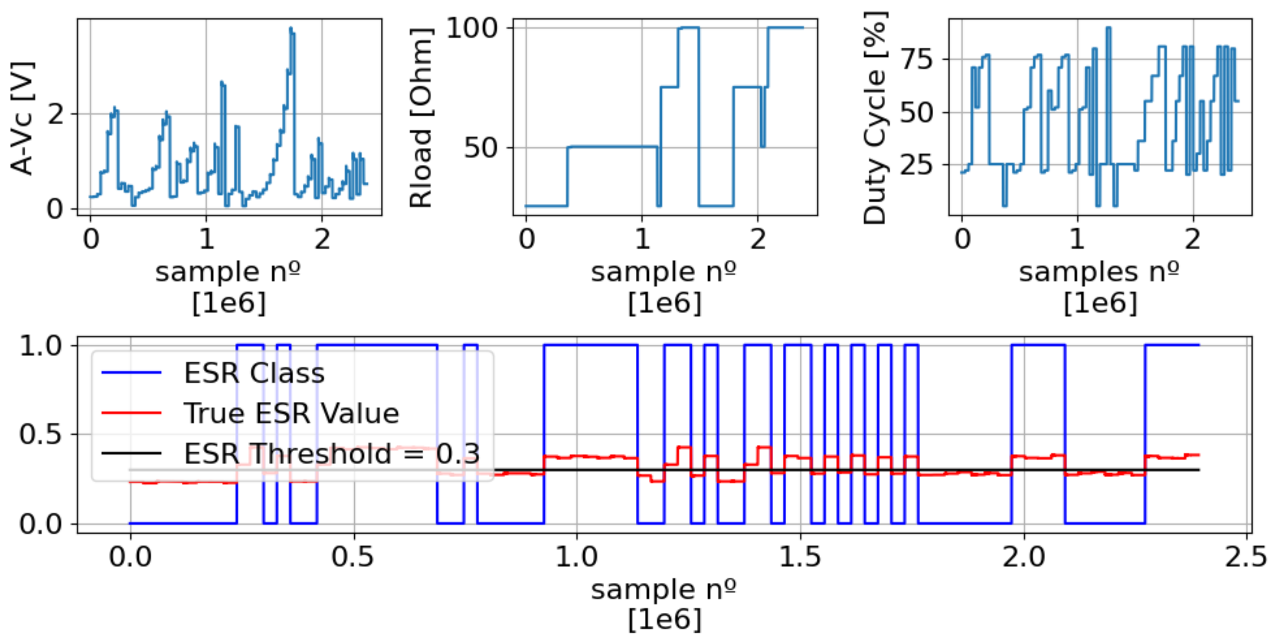

Figure 14. New dataset.

Figure 14 contains 80 new simulations that present completely different operating conditions from those shown in Figure 4.

Regarding the LR model, its response is particularly inaccurate as can be seen in the following figure (Figure 15).

The previous figure shows an accuracy and precision of 55%, and a recall of 46%, which reiterates the analysis carried out in Sections 6 and 7: the LR model should not be used as it is quite imprecise.

Next, it is presented the DTC model response (Figure 16).

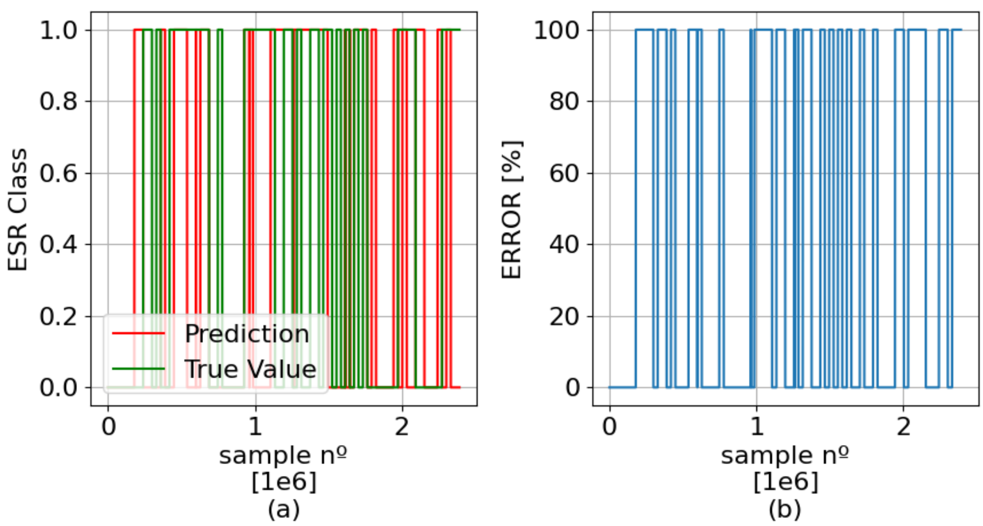

Figure 15. LR model predictions (dataset of Figure 14): (**a**) ESR Class; (**b**) Error.

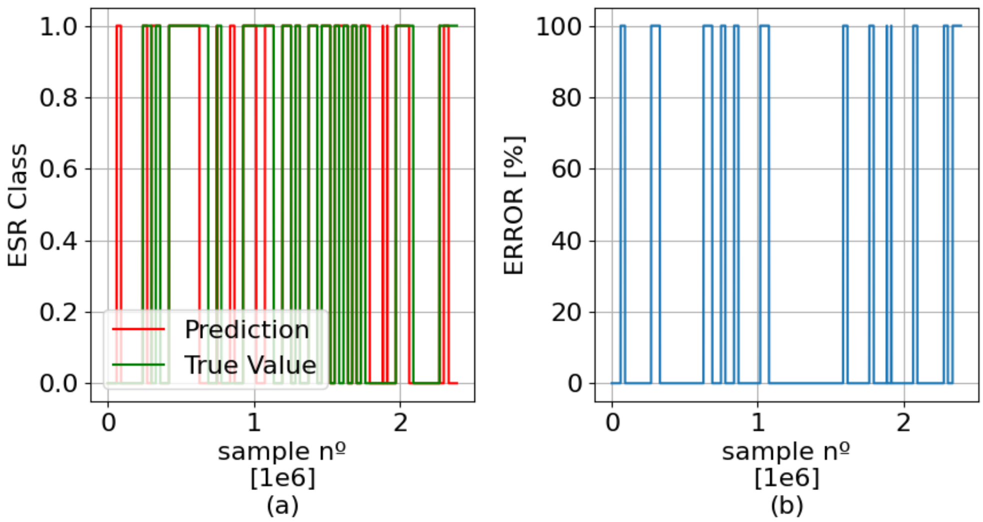

Figure 16. DTC model predictions (dataset of Figure 14): (**a**) ESR Class; (**b**) Error.

Regarding the DTC model, it presents a more precise response when compared to the LR model, with an accuracy of 81%, a precision of 85%, and a recall of 75%. However, the response does not seem to be the most suitable, particularly with regard to the recall.

Attention is drawn to the fact that for the problem under analysis recall is the most important metric, as it is essential that the model correctly identifies the scenarios in which a failure effectively occurred.

Following this, the KNNC model response is presented (Figure 17).

The KNNC model presents a more precise response than the two previous models, with an accuracy of 90%, a precision of 89%, and a recall of 90%.

However, the model that presents the best response is the RFC with an accuracy, precision and recall of approximately 100% as can be seen in the following figure (Figure 18).

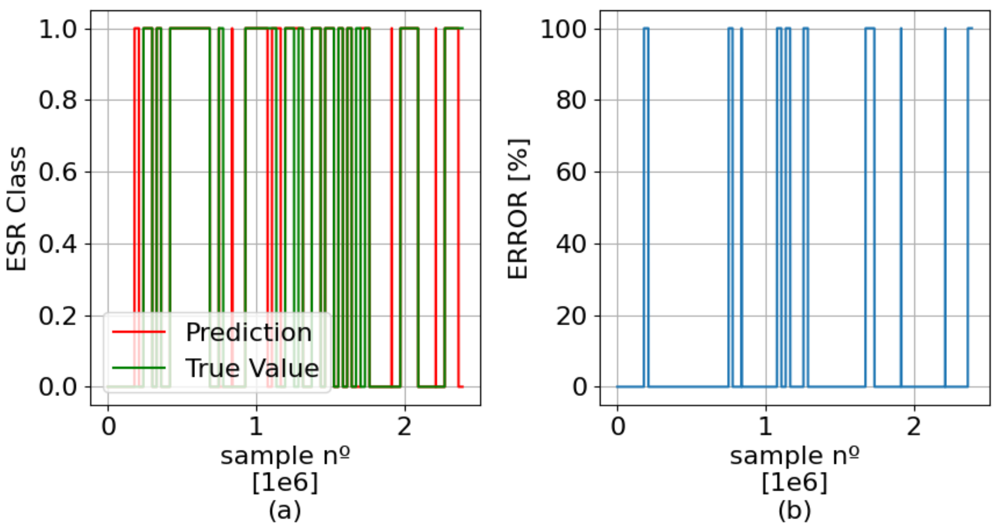

Figure 17. KNNC model predictions (dataset of Figure 14): (**a**) ESR Class; (**b**) Error.

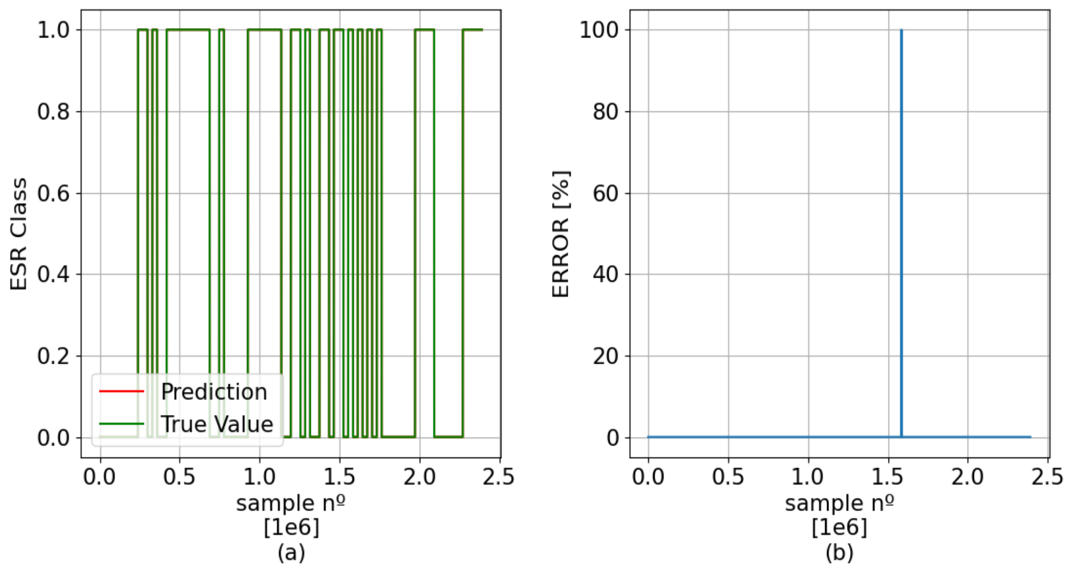

Figure 18. RFC model predictions (dataset of Figure 14): (**a**) ESR Class; (**b**) Error.

9. Conclusions

Our society is still very much dependent on fossil fuels, which, in addition to being non-renewable, pollute the atmosphere and contribute to global warming. In this context, renewable power generation systems (RPGS) play a key role. Therefore, it is crucial to increase the reliability of these systems so that the maximum possible benefit can be achieved. Power converters are one fundamental element of RPGS, with aluminum electrolytic capacitors (AECs) being one of its components. These capacitors are responsible for more than 25% of converter failures; therefore, it is essential to ensure their constant monitoring in order to assess their health condition. Hence, it is possible to implement preventive or predictive maintenance strategies that significantly reduce the unexpected stoppage of RPGS.

This paper presents a non-invasive fault diagnosis technique (NIFDT) for AECs that combines the short-time least square Prony's (STLSP) algorithm with a machine learning (ML) model. The STLSP algorithm processes one of the main attributes of the ML model, the converter output voltage amplitude at the converter switching frequency. The other attributes produced by the STLSP model were evaluated using unsupervised feature selection methods, which proved to be irrelevant. The remaining attributes used by the ML model are the duty cycle and the load resistance, whose measurement does not require the use of sensors inside the converter. In this way, one of the main constraints presented by the vast majority of existing AEC fault diagnosis techniques is overcome.

In this paper, four ML algorithms were evaluated, and it was demonstrated that the random forest classifier is the algorithm that best fits the proposed diagnostic technique.

NIFDT design for AECs has attracted great attention compared to invasive approaches, as they manage to overcome the practical constraints associated with the introduction of sensors inside converters. Hence, the evaluation of the suitability of the proposed solution in the context of other DC link applications, such as PFC, ASD, and PV inverters, will be assessed in future work. Another relevant aspect is the evaluation of the applicability of the proposed solution in more complex systems, such as DC microgrids, and for this purpose, real-time simulators will be used.

Author Contributions: Conceptualization, A.M.R.A., K.L. and A.J.M.C.; methodology, A.M.R.A., K.L. and A.J.M.C.; software, A.M.R.A. and K.L.; validation, A.M.R.A. and K.L.; formal analysis, A.M.R.A. and K.L.; investigation, A.M.R.A., K.L. and A.J.M.C.; resources, A.M.R.A. and A.J.M.C.; data curation, A.M.R.A. and K.L.; writing—original draft preparation, A.M.R.A. and K.L.; writing—review and editing, A.M.R.A. and A.J.M.C.; visualization, A.M.R.A., K.L. and A.J.M.C.; supervision, A.M.R.A.

and A.J.M.C.; project administration, A.M.R.A. and A.J.M.C.; funding acquisition, A.M.R.A. and A.J.M.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Portuguese Foundation for Science and Technology (FCT) under Projects UIDB/04131/2020 and UIDP/04131/2020.

Data Availability Statement: The study did not report any data.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wang, H.; Blaabjerg, F. Reliability of capacitors for DC-link applications in power electronic converters—An overview. *IEEE Trans. Ind. Appl.* **2014**, *50*, 3569–3578. [CrossRef]
2. Lu, W.; Lu, X.; Han, J.; Zhao, Z.; Du, X. Online estimation of ESR for DC-link capacitor of boost PFC converter using wavelet transform based time–frequency analysis method. *IEEE Trans. Power Electron.* **2019**, *35*, 7755–7764. [CrossRef]
3. Wu, Y.; Du, X. A VEN condition monitoring method of DC-link capacitors for power converters. *IEEE Trans. Ind. Electron.* **2018**, *66*, 1296–1306. [CrossRef]
4. Ahmad, M.W.; Kumar, P.N.; Arya, A.; Anand, S. Noninvasive technique for DC-link capacitance estimation in single-phase inverters. *IEEE Trans. Power Electron.* **2017**, *33*, 3693–3696. [CrossRef]
5. Amaral, A.M.R.; Cardoso, A.J.M. Simple experimental techniques to characterize capacitors in a wide range of frequencies and temperatures. *IEEE Trans. Instrum. Meas.* **2010**, *59*, 1258–1267. [CrossRef]
6. Sundararajan, P.; Sathik, M.H.M.; Sasongko, F.; Tan, C.S.; Tariq, M.; Simanjorang, R. Online condition monitoring system for DC-link capacitor in industrial power converters. *IEEE Trans. Ind. Appl.* **2018**, *54*, 4775–4785. [CrossRef]
7. Amaral, A.M.R.; Cardoso, A.J.M. On-line fault detection of aluminium electrolytic capacitors, in step-down DC–DC converters, using input current and output voltage ripple. *IET Power Electron.* **2012**, *5*, 315–322. [CrossRef]
8. Laadjal, K.; Sahraoui, M.; Cardoso, A.J.M.; Amaral, A.M.R. Online estimation of aluminum electrolytic-capacitor parameters using a modified prony's method. *IEEE Trans. Ind. Appl.* **2018**, *54*, 4764–4774. [CrossRef]
9. Cardoso, A.J.M. *Diagnosis and Fault Tolerance of Electrical Machines and Power Electronics*; Institution of Engineering and Technology: London, UK, 2018.
10. Laadjal, K.; Sahraoui, M.; Cardoso, A.J.M. On-line fault diagnosis of DC-link electrolytic capacitors in boost converters using the STFT technique. *IEEE Trans. Power Electron.* **2020**, *36*, 6303–6312. [CrossRef]
11. Zhao, Z.; Davari, P.; Lu, W.; Wang, H.; Blaabjerg, F. An overview of condition monitoring techniques for capacitors in DC-link applications. *IEEE Trans. Power Electron.* **2020**, *36*, 3692–3716. [CrossRef]
12. Dang, H.-L.; Kwak, S. Review of health monitoring techniques for capacitors used in power electronics converters. *Sensors* **2020**, *20*, 3740. [CrossRef] [PubMed]
13. Ren, L.; Zhang, L.; Gong, C. ESR Estimation Schemes of Output Capacitor for Buck Converter from Capacitor Perspective. *Electronics* **2020**, *9*, 1596. [CrossRef]
14. Laadjal, K.; Bento, F.; Cardoso, A.J.M. On-Line Diagnostics of Electrolytic Capacitors in Fault-Tolerant LED Lighting Systems. *Electronics* **2022**, *11*, 1444. [CrossRef]
15. Li, C.; Yu, Y.; Yang, Z.; Liu, Q.; Peng, X. ESR estimation for aluminum electrolytic capacitor of power electronic converter based on compressed sensing and wavelet transform. *IEEE Trans. Ind. Electron.* **2021**, *69*, 1948–1957. [CrossRef]
16. Zhang, Y.; Lian, Z.; Fu, W.; Chen, X. An ESR quasi-online identification method for the fractional-order capacitor of forward converters based on variational mode decomposition. *IEEE Trans. Power Electron.* **2021**, *37*, 3685–3690. [CrossRef]
17. Agarwal, N.; Ahmad, M.W.; Anand, S. Quasi-online technique for health monitoring of capacitor in single-phase solar inverter. *IEEE Trans. Power Electron.* **2017**, *33*, 5283–5291. [CrossRef]
18. Miao, W.; Liu, X.; Lam, K.H.; Pong, P.W.T. Condition monitoring of electrolytic capacitors in boost converters by magnetic sensors. *IEEE Sens. J.* **2019**, *19*, 10393–10402. [CrossRef]
19. Gupta, Y.; Ahmad, M.W.; Narale, S.; Anand, S. Health estimation of individual capacitors in a bank with reduced sensor requirements. *IEEE Trans. Ind. Electron.* **2018**, *66*, 7250–7259. [CrossRef]
20. Agarwal, N.; Arya, A.; Ahmad, M.W.; Anand, S. Lifetime monitoring of electrolytic capacitor to maximize earnings from grid-feeding PV system. *IEEE Trans. Ind. Electron.* **2016**, *63*, 7049–7058. [CrossRef]
21. Ahmad, M.W.; Agarwal, N.; Anand, S. Online monitoring technique for aluminum electrolytic capacitor in solar PV-based DC system. *IEEE Trans. Ind. Electron.* **2016**, *63*, 7059–7066. [CrossRef]
22. Ghadrdan, M.; Peyghami, S.; Mokhtari, H.; Blaabjerg, F. Condition monitoring of dc-link electrolytic capacitor in back-to-back converters based on dissipation factor. *IEEE Trans. Power Electron.* **2022**, *37*, 9733–9744. [CrossRef]
23. Li, H.; Xiang, D.; Han, X.; Zhong, X.; Yang, X. High-accuracy capacitance monitoring of DC-link capacitor in VSI systems by LC resonance. *IEEE Trans. Power Electron.* **2019**, *34*, 12200–12211. [CrossRef]
24. Wang, H.; Wang, H.; Wang, Z.; Zhang, Y.; Pei, X.; Kang, Y. Condition monitoring for submodule capacitors in modular multilevel converters. *IEEE Trans. Power Electron.* **2019**, *34*, 10403–10407. [CrossRef]

25. Dang, H.-L.; Kwak, S.; Choi, S. Identifying DC series and parallel arcs based on deep learning algorithms. *IEEE Access* **2022**, *10*, 76386–76400. [CrossRef]
26. Park, C.-J.; Dang, H.-L.; Kwak, S.; Choi, S. Detection algorithms of parallel arc fault on AC power lines based on deep learning techniques. *J. Electr. Eng. Technol.* **2022**, *17*, 1195–1205. [CrossRef]
27. Dang, H.-L.; Kim, J.; Kwak, S.; Choi, S. Series DC arc fault detection using machine learning algorithms. *IEEE Access* **2021**, *9*, 133346–133364. [CrossRef]
28. Dang, H.-L.; Kwak, S.; Choi, S. Parallel DC arc failure detecting methods based on artificial intelligent techniques. *IEEE Access* **2022**, *10*, 26058–26067. [CrossRef]
29. Dang, H.-L.; Kwak, S.; Choi, S. Different domains based machine and deep learning diagnosis for DC series arc failure. *IEEE Access* **2021**, *9*, 166249–166261. [CrossRef]
30. Park, C.-J.; Dang, H.-L.; Kwak, S.; Choi, S. Deep learning-based series AC arc detection algorithms. *J. Power Electron.* **2021**, *21*, 1621–1631. [CrossRef]
31. Sun, Z.; Bae, S. Multiple-Input Soft-Switching DC–DC Converter to Connect Renewable Energy Sources in a DC Microgrid. *IEEE Access* **2022**, *10*, 128380–128391. [CrossRef]
32. Revathi, B.S.; Prabhakar, M. Solar PV Fed DC Microgrid: Applications, Converter Selection, Design and Testing. *IEEE Access* **2022**, *10*, 87227–87240. [CrossRef]
33. CDM Cornell Dubilier. Aluminium Electrolytic Capacitor Application Guide, Application Note. 2003. Available online: https://www.cde.com/resources/technical-papers/AEappGuide.pdf (accessed on 16 February 2023).
34. Amaral, A.M.R.; Cardoso, A.J.M. Simulation Tool to Evaluate Fault Diagnosis Techniques for DC-DC Converters. *Symmetry* **2022**, *14*, 1886. [CrossRef]
35. Yao, K.; Tang, W.; Hu, W.; Lyu, J. A current-sensorless online ESR and C identification method for output capacitor of buck converter. *IEEE Trans. Power Electron.* **2014**, *30*, 6993–7005. [CrossRef]
36. Senthilkumar, V.; Kumar, B.V. A Survey on Feature Selection Method for Product Review. In Proceedings of the 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), Coimbatore, India, 8–9 October 2021; pp. 1–5.
37. Pudjihartono, N.; Fadason, T.; Kempa-Liehr, A.W.; O'Sullivan, J.M. A review of feature selection methods for machine learning-based disease risk prediction. *Front. Bioinform.* **2022**, *2*, 927312. [CrossRef] [PubMed]
38. Alsolai, H.; Roper, M. A systematic review of feature selection techniques in software quality prediction. In Proceedings of the 2019 International Conference on Electrical and Computing Technologies and Applications (ICECTA), Ras Al Khaimah, United Arab Emirates, 19–21 November 2019; pp. 1–5.