

Article

Data-Driven Low-Frequency Oscillation Event Detection Strategy for Railway Electrification Networks

David Gonzalez-Jimenez , Jon Del-Olmo , Javier Poza , Fernando Garramiola  and Patxi Madina 

Faculty of Engineering, Mondragon Unibertsitatea, 20500 Arrasate-Mondragon, Spain

* Correspondence: dgonzalez@mondragon.edu; Tel.: +34-943794700

Abstract: Low-frequency oscillations (LFO) occur in railway electrification systems due to the incorporation of new trains with switching converters. As a result, the increased harmonic content can cause catenary stability problems under certain conditions. Most of the research published on this topic to date is focused on modelling the event and analysing it using frequency spectrums. However, in recent years, due to the new technologies linked to Big Data (BD) and data mining (DM), a new opportunity to study and detect LFO events by means of machine-learning (ML) methods has emerged. Trains continuously collect data from the most important catenary variables, which offers new resources for analysing this type of event. Therefore, this article presents the design and implementation of a data-driven LFO event detection strategy for AC railway network scenarios. Compared to previous investigations, a new approach to analyse and detect LFO events, based on field data and ML, is presented. To obtain the most appropriate detection approach for the context of this application, on the one hand, this investigation includes a comparison of machine-learning algorithms (support vector machine, logistic regression, random forest, k-nearest neighbours, naïve Bayes) which have been trained with real field data. On the other hand, an analysis of key parameters and features to optimize event detection is also included. Thus, the most significant result of this work is the high metric values of the solution, reaching values above 97% in accuracy and 93% in F-1 score with the random forest algorithm. In addition, the applicability and training of data-driven methods with real field data are demonstrated. This automatic detection strategy can help with speeding up and improving LFO detection tasks that used to be performed manually. Finally, it is worth mentioning that this research has been structured based on the CRISP-DM methodology, established as the de facto approach for industrial DM projects.

Keywords: fault detection; low-frequency oscillation; railway catenary; CRISP-DM; data mining; machine learning; fault diagnosis; field dataset; data-driven



Citation: Gonzalez-Jimenez, D.; Del-Olmo, J.; Poza, J.; Garramiola, F.; Madina, P. Data-Driven Low-Frequency Oscillation Event Detection Strategy for Railway Electrification Networks. *Sensors* **2023**, *23*, 254. <https://doi.org/10.3390/s23010254>

Academic Editors: Francesc Pozo, Mohammad N Noori, Jose Alfonso Antonino-Daviu and Steven Chatterton

Received: 24 November 2022

Revised: 20 December 2022

Accepted: 23 December 2022

Published: 26 December 2022



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The integration of modern power electronics systems is causing power quality and stability problems in electric grids around different sectors. This problem has emerged, for example, in wind energy generation [1] and applications with modular multi-level converters for HVDC [2]. Moreover, the railway sector is another example where similar problems have arisen due to the incorporation of trains with new converters [3].

As the years have passed, the standard EN50388-2012 that railway manufacturers must comply with has become more and more demanding trying to avoid these AC catenary problems [4]. This standard defines different types of instability events based on the harmonic content of catenary currents and voltages. However, an exhaustive analysis of this phenomenon is difficult since the conditions under which it occurs are very specific and difficult to replicate in the laboratory. This is because the occurrence of the events depends on the characteristics of the catenary, on the number of trains connected simultaneously to the same catenary sector, as well as on the type of consumption. This is why it has been hard to propose methods that are able to detect catenary instability problems automatically.

However, for some years now, rolling stock has been equipped with remote monitoring systems that allow detailed analysis of systems operation almost in real time. Having acquired dynamic time series and environmental operational constraints (EOC), data from each train open up new possibilities for the detection and even diagnosis of catenary oscillation events. This context presents an opportunity for the development of data-driven strategies for automatic fault detection and diagnostics. Although data are available, Big Data and cloud-based monitoring platforms do not have enough acquisition frequency, for now, in order to analyse the harmonic content of catenary voltage and current as proposed by the standards. Therefore, another detection approach is needed, being that data-driven detection approaches based on machine learning are more suitable for the application.

This scientific paper presents the design and implementation of an ML-based strategy for the detection of LFO events in AC railway network scenarios. In order to achieve the main objective and obtain the most suitable detection approach for the application under study, on the one hand, this investigation includes a comparison of machine-learning algorithms (support vector machine, logistic regression, random forest, k-nearest neighbours, and naïve Bayes) which have been trained with real field data, and on the other hand, an analysis of key parameters and features to optimize event detection. In summary, the scientific contributions of this research are the following:

1. An automatic fault detection strategy based on ML is designed and implemented. Most of the studies carried out to date have focused only on the LFO event modelling and analysis of the possible causes [5–12]. The authors in [7] also analyse the influence that the characteristics of the traction units have on the occurrence of oscillations. This could lead to hardware and control software design improvements to mitigate these events.
2. The dataset used for both training and testing of the algorithms is based on field data. That is, the industrial partner has detected and recorded manually different LFO events during the operation of its train fleet, which facilitates the development of supervised ML methodologies. This differs from the rest of the articles, as most of them are based on simulation or testbench data.
3. Most DM scientific works follow the same steps to prepare data, train ML models, and evaluate their results. However, they do not follow any industry standard, which could make it difficult to deploy in an industrial application. Therefore, in this research, the DM project is structured based on the CRISP-DM methodology, established as the de facto approach for industrial DM projects [13–15]. Figure 1 shows the CRISP-DM methodology and its main steps.

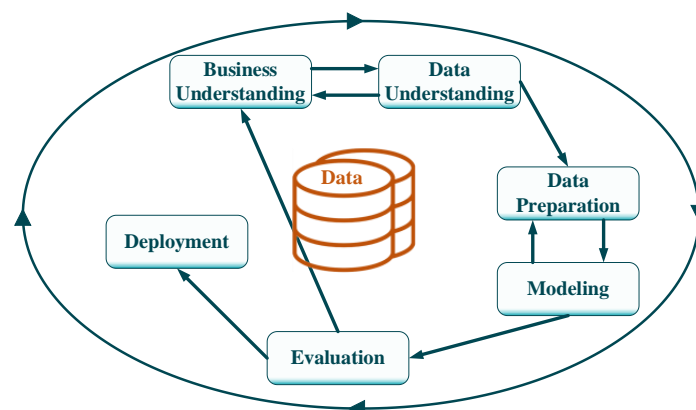


Figure 1. CRISP-DM methodology cycle.

The structure of the article follows the CRISP-DM cycle, as shown in Figure 1. Section 2 is focused on the theoretical analysis of the phenomena under investigation, as well as on the available raw dataset. This is equivalent to the Business Understanding and Data Understanding steps from the aforementioned methodology. Section 3 covers everything

related to data preparation. Section 4 summarizes the different tasks regarding the modelling step, where the different ML algorithms are trained. Finally, Section 5 copes with the deployment, explaining the tools and platforms that have been used.

2. Power Quality and Stability Problems in Railway Applications

As it was mentioned before, catenary instability problems occur due to the incorporation of new trains with modern converters, since these are the largest source of harmonics in the entire system [16,17]. Concerning instabilities, there are three main phenomena that, due to the interaction between the traction power supply network and the trains, jeopardize the quality of supply and the stability of the catenary: the harmonic resonance, the harmonic instability, and the low-frequency oscillation (LFO) [6,18]. The characteristics of each type of event are presented in [18] and are summarized in Figure 2. Formation conditions and characteristics of each phenomenon are shown, as well as the influential factors at train and network levels.

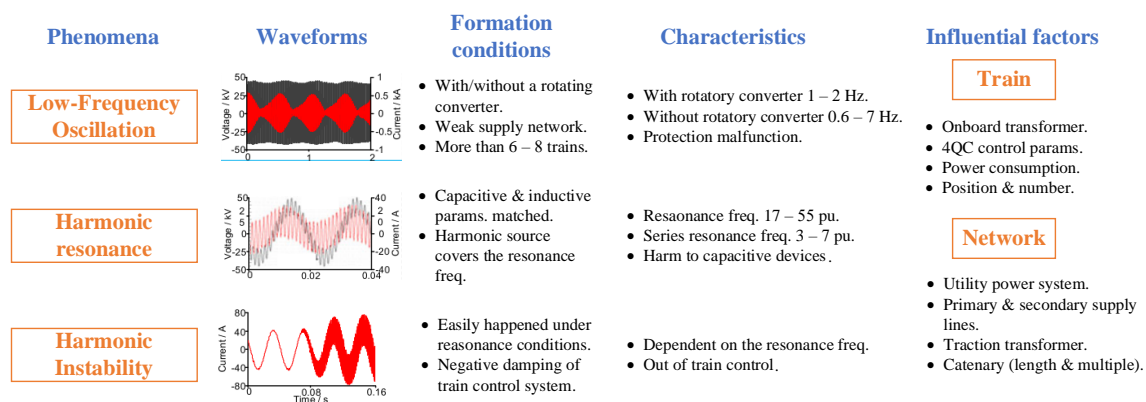


Figure 2. Main characteristics of the most common AC catenary instability phenomena: LFO, harmonic resonance, and harmonic instability [18].

This investigation studies the possibility of applying automatic detection approaches for LFO events, as these have been detected by the industrial partner in greater quantity among the dataset collected from its fleet. LFO is characterized by a low-frequency envelope of catenary voltage and current caused by inter-harmonic components at $f_0 \pm f_{osc}$ around the fundamental frequency (f_0) [9]. As [18] proposes, LFO events are detected in two types of scenarios:

1. In a railway system which has adopted the rotary frequency converter (RFC) as the power supply solution. The typical oscillatory frequency (f_{osc}) is approximately 10–30% of the respective power system's fundamental frequency (f_0).
2. In a railway system equipped with a static frequency converter (SFC) as a power supply solution and where several trains with four quadrant converters are in standby mode (auxiliary loads) located in the same railway depot, away from the traction substation (TSS). This leads to a 0.6–7 Hz LFO of the catenary voltage.

According to [19], the factors that contribute more to the appearance of this kind of event are:

- The number of vehicles and the load current in the same network sector.
- The contact line distance. The longer the contact line is, the larger the source impedance that increases the probability of instabilities will be.
- The line side traction converter control parameter tuning.

Figures 3 and 4 show snapshots (in RMS values) of an LFO event detected by the industrial partner. In this case, the phenomena happened while the train was stopped (in a depot).

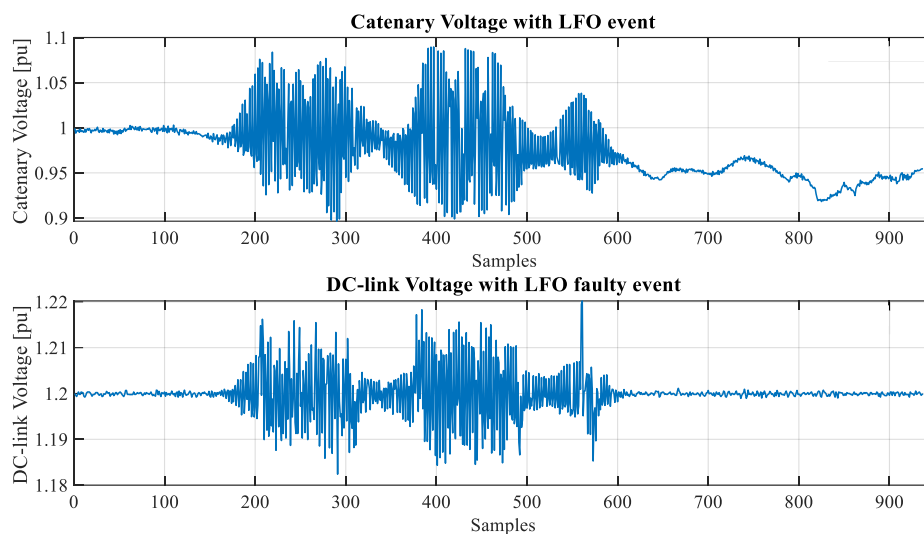


Figure 3. Catenary and DC-Link voltage (pu) under LFO faulty event.

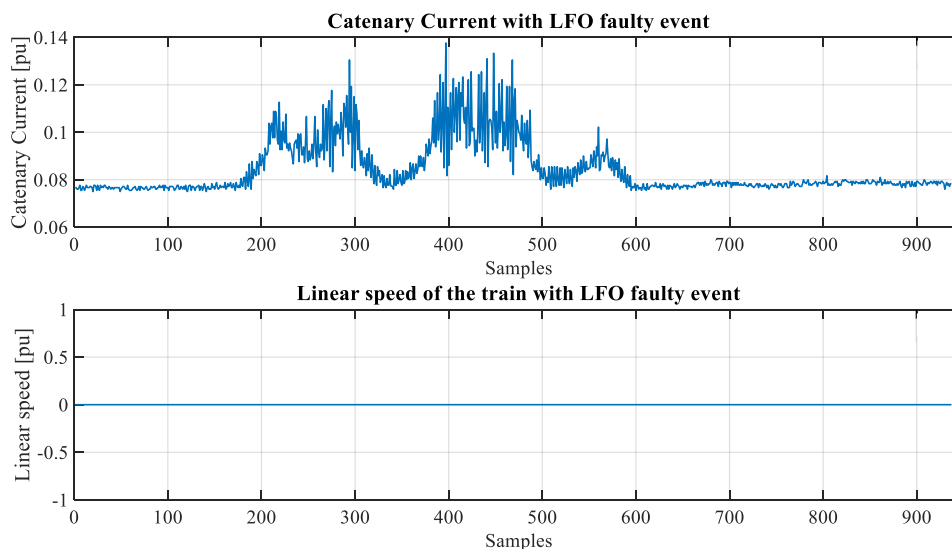


Figure 4. Catenary current and speed (pu) under LFO faulty event.

Most of the research published so far analyses this phenomenon from an analytical point of view (modelling the system) [7,8,10–12,20–22] or by analysing the harmonic content of field data [3,23,24]. EN50388-2012 also proposes the frequency as one of the features to study in order to distinguish between healthy and faulty behaviour. Therefore, once an event occurs, it is relatively easy to detect it by looking for the corresponding frequency components. This is a simple solution that does not require major technical developments. However, frequency analysis requires the analysed signals to have a sampling rate fast enough to distinguish faulty harmonics and avoid aliasing. Although nowadays railway traction control units (TCU) sample catenary voltage and current at high frequencies (e.g., 20 kHz) for drive control purposes, these measurements are usually down-sampled to send them to remote monitoring platforms. Unless there is a specific need, variables acquired at high frequency are not remotely available. As a result, the bottleneck in these cases is the computational and the data storage capacity that monitoring platforms have nowadays.

In this research, the data related to electrical variable measurements are down-sampled to several milliseconds. Moreover, instantaneous measurements are not available since sinusoidal signals were processed and stored as RMS values. Figure 5 shows a detailed view of catenary and DC-Link voltage in per unit with an LFO event.

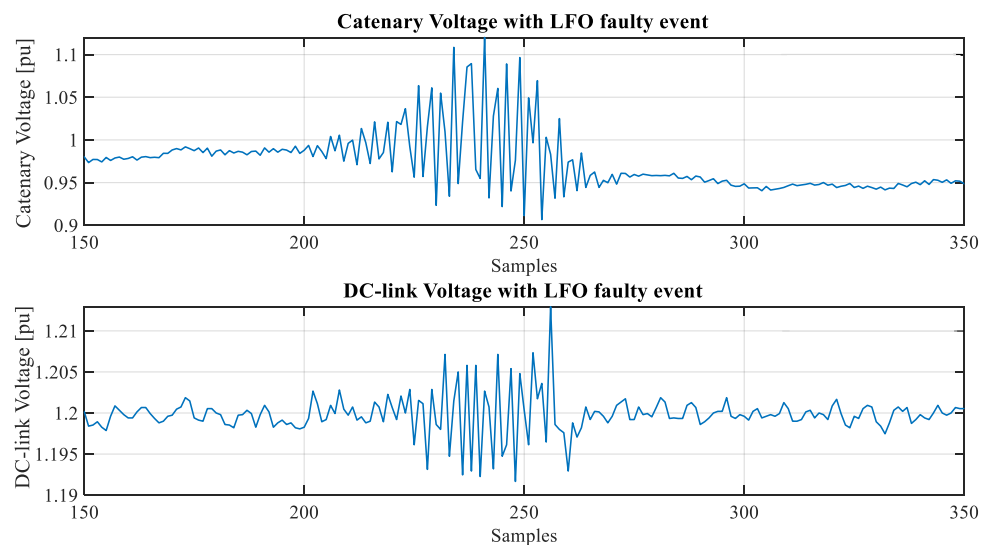


Figure 5. Detail of catenary and DC-Link voltage (pu).

In this context, it is considered that a data-driven detection strategy helps to identify this type of event, even with under-sampled data. Furthermore, the industrial partner has already identified 35 LFO events in the remote monitoring dataset. This is a good initial source of information to start the steps of designing a supervised ML strategy for automatic detection of LFO phenomena.

3. Data Preparation

3.1. Data Selection

In order to create a suitable database for training and testing ML algorithms, the most important variables are selected from all the available ones in the main data transmission bus of the train. So, as the main objective is to deal with LFO (an electrical phenomenon), and the first variable selection round is focused on the electric variables related to the drive of the train and its catenary. Specifically, the following variables are selected with the help of expert knowledge: catenary voltage, DC-bus voltage, catenary current, motor stator currents, estimated power, speed, and torque.

In addition, as a way of putting the results into context, additional EOC variables such as GPS latitude and longitude are included. These variables are intended to improve the explainability of ML algorithms and to help with identifying problematic railway network sections and scenarios.

Finally, the correlation between all the variables is analysed. In this way, redundant variables are eliminated while maintaining the highest possible level of information in the dataset. Thus, if a correlation between two variables is higher than 90 %, only one of them is considered. After analysing the correlation percentages between all the variables, a total of 29 elements are selected as the useful dataset from an initial dataset of 103 variables selected using expert knowledge.

3.2. Data Cleaning

Data cleaning consisted of:

- Standardizing and adapting variable names.
- Filtering null values and replacing them using forward and backward filling techniques.
- Detecting and cleaning outliers.

In the case of outliers, apart from atypical values due to acquisition anomalies, pantograph detachments posed an obstacle when identifying catenary oscillations, because both events differ considerably from normal operation. In order to prevent false alarms and based on expert knowledge, it is determined that catenary voltages above and below

25% with respect to the nominal value are not necessary to detect LFO events, so they are classified as outliers.

3.3. Data Construction

Once the available dataset is cleaned of null values and redundancies, feature engineering is applied. The goal of this step is to prepare a training dataset that best fits the ML algorithms, as well as to improve the performance of the models.

For manual feature extraction, time-domain, frequency-domain, and time/frequency-domain techniques can be used. In this study, considering the sampling rate of the variables, techniques including frequency analysis are discarded. Hence, the approach of statistical time-domain feature extraction is performed. For each variable, the acquired event length is split into windows of N samples. Then, these windows are summarized, computing the statistical characteristics shown in Table 1.

Table 1. Time-domain statistical features extracted from the cleaned dataset.

Statistical Feature	Description
Maximum	$\text{Max}(x_i, \dots, x_N)$
Minimum	$\text{Min}(x_i, \dots, x_N)$
Average	$\frac{1}{N} \sum_{i=1}^N x_i$
Variance	$\frac{\sum_{i=1}^N (x_i - \mu)^2}{N-1}$
Standard Deviation	$\sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N-1}}$
Range	$\text{Max}(x_i, \dots, x_N) - \text{Min}(x_i, \dots, x_N)$
Root Mean Square	$\sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}$
Kurtosis	$\frac{\sum_{i=1}^N (x_i - \mu)^4}{N\sigma^4}$
Skewness	$\frac{1}{N} \frac{\sum_{i=1}^N (x_i - \mu)^3}{\sigma^3}$

As a result, the number of samples for each variable in a log is reduced, and the number of features increased. That is why, for one event log, the initial raw data window of 29 variable (columns) and 937 samples (rows) is transformed into 237 features and 93 samples. Figure 6 shows the time series of a variable (catenary voltage) and some of the extracted statistical features (mean and variance).

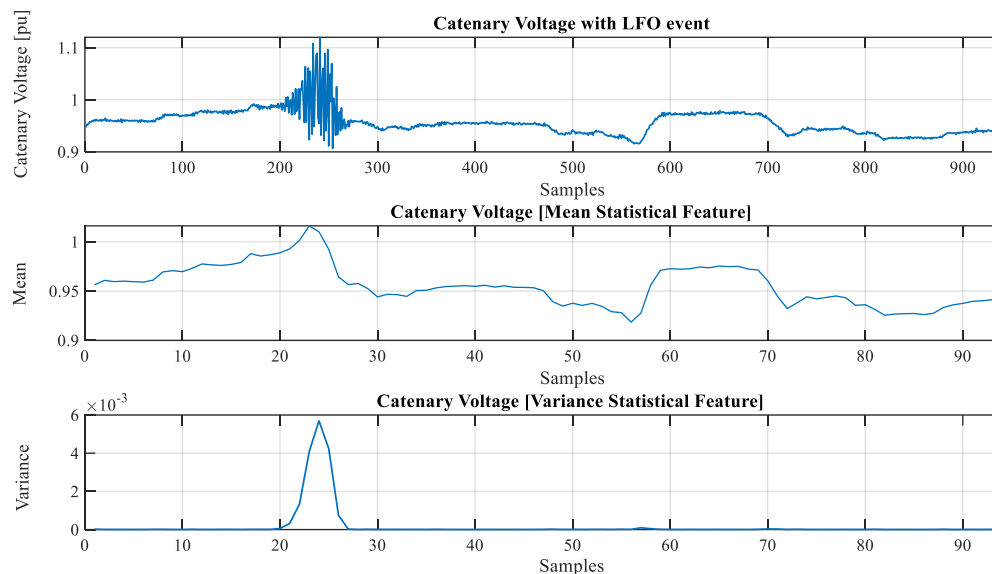


Figure 6. Statistical time-domain features from catenary voltage (mean and variance features).

Even if the dimension of the raw dataset is reduced considerably, feature selection techniques are applied to decrease the dimensionality even more. To do so, an analysis of variance algorithm (F-test ANOVA) is used to rank the independent features depending on their explanatory capacity with respect to the dependent variable (target). In other words, this method provides the N best characteristics that show the best separability between classes of the target variable, based on variance calculations [25,26]. In this case, the target variable is the label that defines if a specific sample is part of an LFO or not. At the end of this process, the dataset dimension is reduced from 237 features to 30, enabling more efficient model training and testing.

Finally, the variables are standardized using a standard scaler, matching them to a standard normal distribution (SND). All the variables are converted to have a mean value of zero and a standard deviation of one. The objective of this operation is to improve algorithm efficiency.

4. Model Training

4.1. Training Phase

Once the dataset is prepared, the machine-learning algorithms should be trained. In a first attempt, it is considered that the basic supervised learning algorithms should be able to detect LFO events with the available information. Therefore, the following algorithms are selected to make a comparison between them and select the most appropriate one for this application:

- Logistic regression (LR).
- Support vector machines (SVM).
- Random forest (RF).
- Naïve Bayes (NB).
- k -Nearest neighbours (k NN).

The next step is to split the whole pre-processed dataset into two parts: 70% for training and 30% for testing. Furthermore, when training and optimizing the selected algorithms, the k -fold cross-validation (CV) technique is applied. This technique consists of randomly dividing the training dataset into k subsets, of which $k-1$ groups is used to train the algorithm, and the remaining group is used for validation (see Figure 7). This iterative process is repeated k times using a different subset as validation in each iteration. Therefore, the process calculates k classification metrics whose average is used as the result in the training.

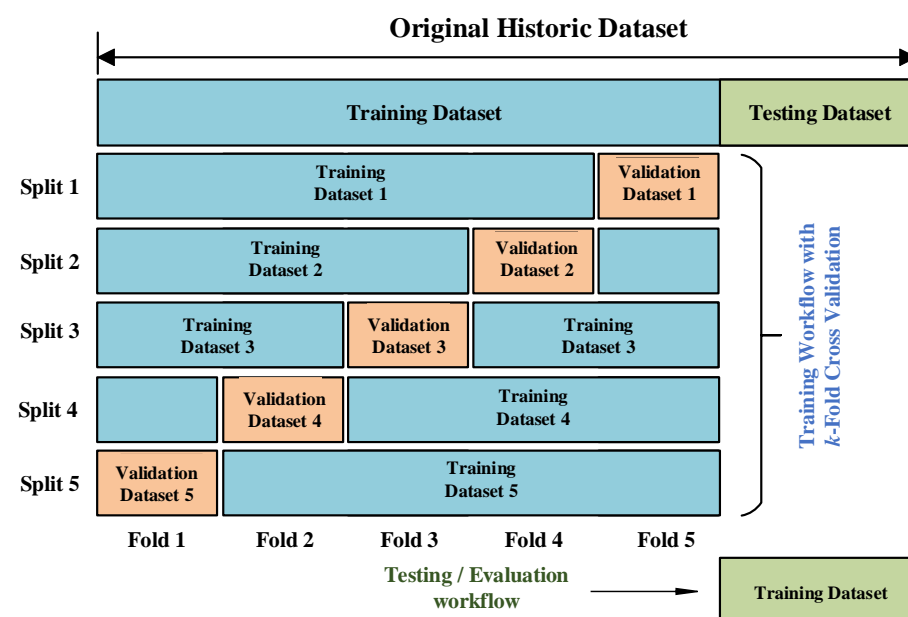


Figure 7. k -Fold cross-validation process.

The main purpose of CV is to protect these kinds of algorithms against overfitting, which are very prone to it. A model with overfitting is highly tuned for the training dataset and is able to predict these data very accurately. However, it fails to predict new data.

In order to optimize the algorithms, hyperparameters are tuned using the Grid Search function from Python's SKLearn package. This method together with the CV process returns the optimal set of hyperparameters with respect to the accuracy loss function.

4.2. Model Assessment

Once different algorithms are trained, accuracy and F-1 score are used for the assessment of the models [27]. Table 2 shows the results.

Table 2. Evaluation metrics of the algorithms after testing.

	LR	SVM	RF	NB	kNN
Accuracy (%)	96.7	97.1	97.1	95.3	97.1
F-1 Score (%)	92.7	93.7	93.9	89.5	93.8

In all the cases, the results show that in terms of accuracy and F-1 score, algorithms are nearly above 90%. Although the algorithm with the best scores is RF, further analysis is performed.

As it was shown before, new features are extracted via the 10-sample time-sliding window method, and the most 30 influential features are selected by the ANOVA F-test. As these two factors are randomly selected, several tests are performed for training the algorithms with different combinations of sliding window length and feature quantity. Knowing that in the future this algorithm may be applied to an entire fleet with dozens of trains and that the dimension of the input dataset will be key, the objective of this analysis is to see the influence of these parameters in the performance of the algorithms. In particular, the window lengths of 5, 10, and 20 samples and datasets with 5, 15, and 30 features are tested (see Table 3). Figures 8 and 9 show the results of each algorithm for different setups.

Table 3. Dataset pre-processing setup description (S1–S9).

	S1	S2	S3	S4	S5	S6	S7	S8	S9
N° of samples/window	10	5	20	10	5	20	10	5	20
N° of selected features	30	30	30	15	15	15	5	5	5

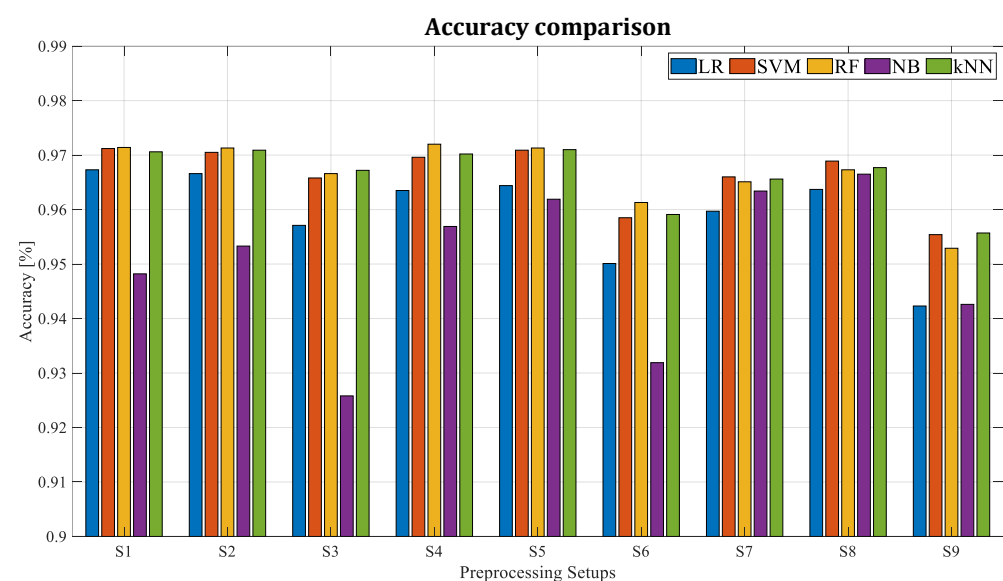


Figure 8. Accuracy for different algorithms and dataset configurations.

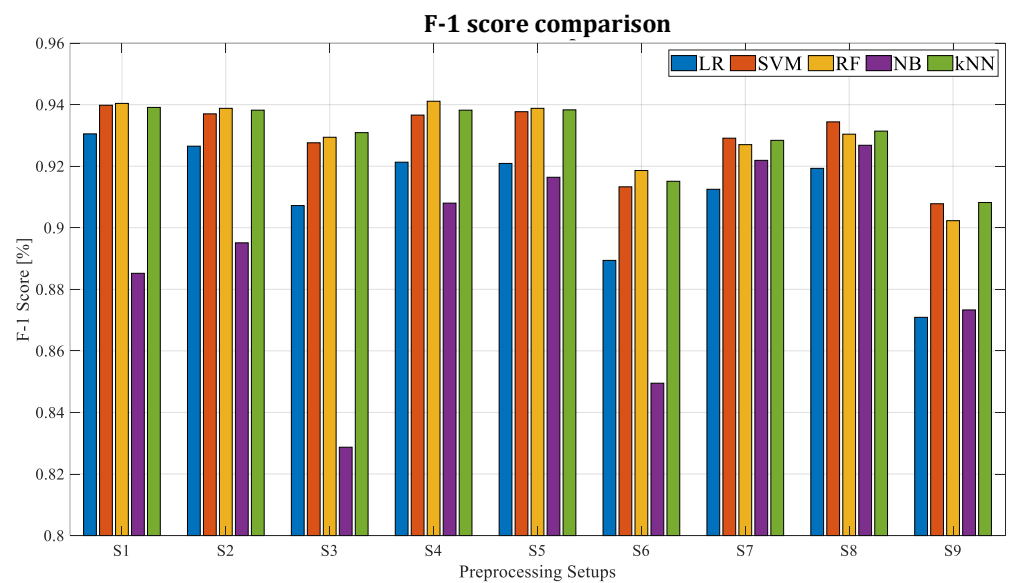


Figure 9. F-1 score for different algorithms and dataset configurations.

All the scores with the different pre-processing configurations improve 90% of accuracy. However, the setups where the window length for feature extraction is composed of 20 samples (Set 3, 6, and 9) are the worst setups. This is because by trying to compress the information of N points into just one, some of the information is lost unintentionally. Therefore, the higher the value of N , the less information is transmitted to the algorithm (more information is lost). Furthermore, five algorithms and pre-processing setups have achieved accuracy values around 97% and high values in F-1 score (RF with Set 1, Set 4, and Set 5, k NN Set 5 and SVM Set 8). It is worth mentioning that Set 4 with random forest uses half as many features as Set 1, obtaining very similar results. This may be a determinant algorithm and setup selection factor during the industrial deployment phase in the future.

5. Deployment

Although the results shown above are a proof of concept, the development is performed using tools used by the industry. Data storage, pre-processing, and modelling are performed using the Amazon Web Services platform. As an example, AWS S3 is used as the main data storage service. In addition, Cloud9 and Amazon SageMaker are used for data pre-processing and algorithm modelling and testing. Moreover, Figure 10 shows a schematic of the developed workflow in AWS. As a result, using those structures in the future for DM projects will reduce the whole detection strategy deployment time.

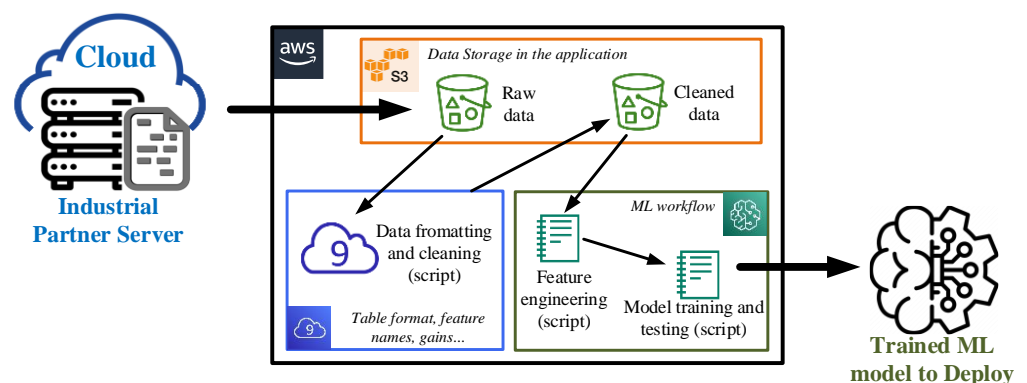


Figure 10. Data-mining architecture in Amazon Web Services platform.

In addition, a final test is performed simulating the conditions that the application will encounter in the real operation. Therefore, instead of feeding the algorithms with data

snapshots of limited duration (minutes), new time series with an approximate duration of eight hours are introduced in the detection workflow. Knowing that in a real environment the detection strategy will have to analyse massive data, the objective is to see its efficiency in a more real environment.

Figure 11 shows the pre-processing and classification times of the five different setups. In general, accuracy and F-1 score maintain the values shown in Table 2, but RF outperforms the rest of the algorithms in terms of pre-processing and prediction times. On average, it takes 94% less time than *k*NN or SVM algorithms.

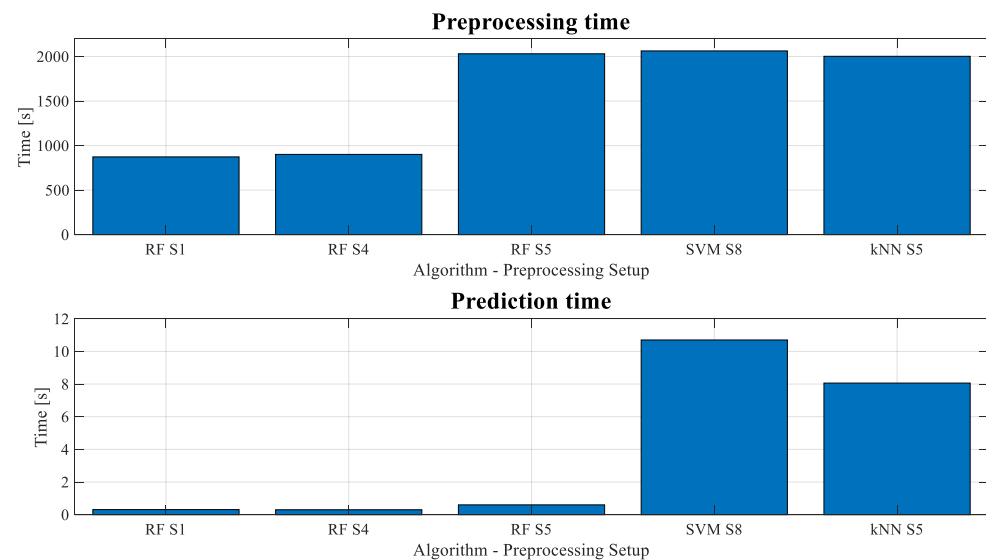


Figure 11. Pre-processing and classification times of the different setups.

Finally, it is worth mentioning that the manual method used previously by the industrial partner to detect LFO events visually is compared to the new proposed strategy. Considering that in the worst-case scenario the shortest LFO event found previously lasted less than 1 minute, the analyst should check one by one 1-minute windows in an 8-hour log and see whether there is an LFO or not. This means analysing 480 windows visually. With an average time of 20 seconds per window, the analysis takes 1 hour and 36 minutes. This is the case where the information is analysed blindly, without further details of date or location. In contrast, automatic detection with the implemented strategy takes approximately 10 minutes, both for the pre-processing and the detection. Moreover, in the 8-hour log test, the automatic strategy found 6% more events compared to the previous manual analysis.

6. Conclusions

This paper presents a data-driven detection strategy for low-frequency oscillation events in AC railway catenaries.

In contrast to previous publications, the main contribution of this article is the development of the automatic detection strategy of the anomalous events based on machine-learning algorithms, which is not found in the scientific literature. In addition, another important contribution is the use of real field data for the training and testing of the ML algorithms. This work shows that data availability strengthens and supports maintenance of industrial applications.

Moreover, the CRISP-DM methodology is used in the development of the proposed solution, which is nowadays considered as the de facto approach in industrial data-mining projects. This has helped to understand and pre-process the historical dataset and in turn to train and evaluate the implemented classification algorithms.

Due to the monitoring design already implemented in the train fleet, the acquisition frequency of the variables is not fast enough to analyse LFO events by means of frequency analysis, as required by EN50388-2012. However, the large dataset available is optimized by

extracting new statistical features in the time domain and filtering redundancies between them by applying feature selection techniques. This allows the use of machine-learning techniques. Specifically, five different supervised classification algorithms (LR, RF, SVM, kNN, and NB) are trained using the fivefold cross-validation approach.

The results in terms of accuracy and F-1 score show that the most suitable classification algorithm is random forest, reaching 97.1% of accuracy and 93.9% of F-1 score. Moreover, it is the fastest classification strategy analysed, considering pre-processing and prediction times.

The general conclusion of the research is that, although the current monitoring does not meet the requirements to analyse LFO failure according to the regulation, the acquisition of a large number of variables, data, and casuistry (Big Data) opens the opportunity to deploy ML models that are proven to be effective and efficient in this application.

Author Contributions: Conceptualization, D.G.-J. and J.D.-O.; investigation, D.G.-J. and J.D.-O.; methodology, D.G.-J. and J.D.-O.; formal analysis, J.P., P.M. and F.G.; project administration, J.P.; validation, D.G.-J. and J.D.-O.; writing—original draft, D.G.-J. and J.D.-O.; writing—review and editing, J.D.-O., J.P., P.M. and F.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

Abbreviations

The following abbreviations are used in this manuscript:

ML	Machine Learning
CRISP-DM	Cross-Industry Standard Process for Data Mining
LFO	Low-Frequency Oscillation
TCU	Traction Control Unit
SND	Standard Normal Distribution
SVM	Support Vector Machine
RF	Random Forest
kNN	k Nearest Neighbour
NB	Naïve Bayes
AWS	Amazon Web Services

References

1. Ebrahimzadeh, E.; Blaabjerg, F.; Wang, X.; Bak, C.L. Modeling and Identification of Harmonic Instability Problems in Wind Farms. In Proceedings of the 2016 IEEE Energy Conversion Congress and Exposition (ECCE), Milwaukee, WI, USA, 18–22 September 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 1–6.
2. Li, T.; Gole, A.M.; Zhao, C. Harmonic Instability in MMC-HVDC Converters Resulting From Internal Dynamics. *IEEE Trans. Power Deliv.* **2016**, *31*, 1738–1747. [[CrossRef](#)]
3. Hu, H.; Zhou, Y.; Li, X.; Lei, K. Low-Frequency Oscillation in Electric Railway Depot: A Comprehensive Review. *IEEE Trans. Power Electron.* **2021**, *36*, 295–314. [[CrossRef](#)]
4. EN 50388:2012 (EN); European Standards (EN); Railway Applications—Fixed Installations and Rolling Stock—Technical Criteria for the Coordination between Power Supply and Rolling Stock to Achieve Interoperability—Part 2: Stability and Harmonics. Available online: <https://standards.iteh.ai/catalog/standards/clc/fd5d3267-2aed-42f9-add6-d7bfcac78762/en-50388-2012> (accessed on 23 November 2022).
5. Hu, H.; Tao, H.; Blaabjerg, F.; Wang, X.; He, Z.; Gao, S. Train–Network Interactions and Stability Evaluation in High-Speed Railways—Part I: Phenomena and Modeling. *IEEE Trans. Power Electron.* **2018**, *33*, 4627–4642. [[CrossRef](#)]
6. Hu, H.; Tao, H.; Wang, X.; Blaabjerg, F.; He, Z.; Gao, S. Train–Network Interactions and Stability Evaluation in High-Speed Railways—Part II: Influential Factors and Verifications. *IEEE Trans. Power Electron.* **2018**, *33*, 4643–4659. [[CrossRef](#)]
7. Frutos, P.; Ladoux, P.; Roux, N.; Larrazabal, I.; Guerrero, J.M.; Briz, F. Low Frequency Stability of AC Railway Traction Power Systems: Analysis of the Influence of Traction Unit Parameters. *Electronics* **2022**, *11*, 1593. [[CrossRef](#)]

8. Wang, Y.; Li, T.; Liu, J.; Liu, F.; Su, Y.; Zhang, Q.; Zhang, W.; Liu, Z. Modeling and Analysis of Low-Frequency Oscillation for Electrified Railway under Mixed Operation of Passenger and Freight Trains. *Energies* **2022**, *15*, 7544. [[CrossRef](#)]
9. Zhou, Y.; Hu, H.; Yang, X.; Yang, J.; He, Z.; Gao, S. Low Frequency Oscillation Traceability and Suppression in Railway Electrification Systems. *IEEE Trans. Ind. Appl.* **2019**, *55*, 7699–7711. [[CrossRef](#)]
10. Laury, J. Stability of Low-Frequency AC Railways Models and Transient Stability. Ph.D. Thesis, Luleå University of Technology, Luleå, Sweden, 2019.
11. Cheng, S.; Ma, L.; Ge, X.; Peng, L.; Liu, H. Low-Frequency Oscillation Analysis in Train-Traction Power Supply System Using an SISO Voltage Loop Model. *IEEE Trans. Transp. Electrification* **2022**, *8*, 636–648. [[CrossRef](#)]
12. Paul, F.G.; Guerrero, J.M.; Muniategui-Aspiazua, I.; Vicente-Makazaga, I.; Endemano-Isasi, A.; Ortega-Rodriguez, D.; Briz, F. Power-Hardware-in-the-Loop Emulation of the Low-Frequency Oscillation Phenomenon in AC Railway Networks. *IEEE Access* **2022**, *10*, 87374–87386. [[CrossRef](#)]
13. Pete, C.; Julian, C.; Randy, K.; Thomas, K.; Thomas, R.; Colin, S.; Wirth, R. *Crisp-Dm 1.0. Step-by-Step Data Mining Guide*; SPSS Inc.: Chicago, IL, USA, 2000.
14. Martinez-Plumed, F.; Contreras-Ochando, L.; Ferri, C.; Hernandez-Orallo, J.; Kull, M.; Lachiche, N.; Ramirez-Quintana, M.J.; Flach, P. CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories. *IEEE Trans. Knowl. Data Eng.* **2021**, *33*, 3048–3061. [[CrossRef](#)]
15. Saleh, S.J.; Ali, S.Q.; Zeki, A.M. Random Forest vs. SVM vs. KNN in Classifying Smartphone and Smartwatch Sensor Data Using CRISP-DM. In Proceedings of the 2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI), Sakheer, Bahrain, 26 October 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–4.
16. Möllerstedt, E.; Bernhardsson, B. Out of Control Because of Harmonics—an Analysis of the Harmonic Response of an Inverter Locomotive. *IEEE Control Syst.* **2000**, *20*, 70–81. [[CrossRef](#)]
17. Danielsen, S. Electric Traction Power System Stability: Low-Frequency Interaction between Advanced Rail Vehicles and a Rotary Frequency Converter. Ph.D. Thesis, Norwegian University of Science and Technology, Trondheim, Norway, 2010.
18. Hu, H.; Tao, H.; Blaabjerg, F.; Wang, X.; He, Z.; Gao, S. Atlab. *IEEE Trans. Power Electron.* **2018**, *33*, 4627–4642. [[CrossRef](#)]
19. Wang, H.; Mingli, W.; Sun, J. Analysis of Low-Frequency Oscillation in Electric Railways Based on Small-Signal Modeling of Vehicle-Grid System in Dq Frame. *IEEE Trans. Power Electron.* **2015**, *30*, 5318–5330. [[CrossRef](#)]
20. Pan, P.; Hu, H.; Yang, X.; Blaabjerg, F.; Wang, X.; He, Z. Impedance Measurement of Traction Network and Electric Train for Stability Analysis in High-Speed Railways. *IEEE Trans. Power Electron.* **2018**, *33*, 10086–10100. [[CrossRef](#)]
21. Danielsen, S.; Molinas, M.; Toftevaag, T.; Fosso, O.B. Constant Power Load Characteristic’s Influence on the Low-Frequency Interaction between Advanced Electrical Rail Vehicle and Railway Traction Power Supply with Rotary Converters. In Proceedings of the Modern Electric Traction (MET), Gdansk, Poland, 22–23 January 2009.
22. Lv, X.; Wang, X.; Member, S.; Che, Y.; Li, L. Dynamic Phasor Modelling and Low-Frequency Oscillation Analysis of the Single-Phase Vehicle-Grid System. *CSEE J. Power Energy Syst.* **2020**, *1*, 1–13. [[CrossRef](#)]
23. Jiang, X.; Hu, H.; He, Z.; Tao, H.; Qian, Q. Study on Low-Frequency Voltage Fluctuation of Traction Power Supply System Introduced by Multiple Modern Trains. *Electr. Power Syst. Res.* **2017**, *146*, 246–257. [[CrossRef](#)]
24. Bian, D.; Yu, Z.; Shi, D.; Diao, R.; Wang, Z. A Robust Real-Time Low-Frequency Oscillation Detection and Analysis (LFODA) System with Innovative Ensemble Filtering. *CSEE J. Power Energy Syst.* **2020**, *6*, 174–183. [[CrossRef](#)]
25. Perangin-Angin, D.J.; Bachtiar, F.A. Classification of Stress in Office Work Activities Using Extreme Learning Machine Algorithm and One-Way ANOVA F-Test Feature Selection. In Proceedings of the 2021 4th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), Yogyakarta, Indonesia, 16 December 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 503–508.
26. Zhuang, H.; Liu, X.; Wang, H.; Qin, C.; Li, Y.; Li, W.; Shi, Y. Diagnosis of Early Stage Parkinson’s Diseases on Quantitative Susceptibility Mapping Using Complex Network with One-Way ANOVA F-Test Feature Selection. *J. Mech. Med. Biol.* **2021**, *21*, 2140026. [[CrossRef](#)]
27. Hossin, M.; Sulaiman, M.N. A Review on Evaluation Metrics for Data Classification Evaluations. *Int. J. Data Min. Knowl. Manag. Process* **2015**, *5*, 1–11. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.