# Bearings' Vibration Fault Detection via Physics-Informed Deep Learning

**Leonardo Streck Raupp[abc*], Isis Didier Lins[abc], Caio Souto Maior[abc], Thiago Cavalcanti[d], Márcio José das Chagas Moura[abc], Gustavo de Novaes Pires Leite[e]**

[a]Center for Studies and Trials in Risk and Ambiental Modelling (CEERMA), Recife, Brazil
[b]Post-Graduate Program in Production Engineering (PPGEP – UFPE), Recife, Brazil
[c]Federal University of Pernambuco (UFPE), Recife, Brazil
[d]University of Pernambuco (UPE), Recife, Brazil
[e]Federal Insitute of Pernambuco (UFPE), Recife, Brazil

**Abstract:** Reliability engineering is crucial for the success and competitiveness of industries heavily reliant on complex equipment. In this context, Condition-Based Maintenance (CBM) and Prognostics and Health Management (PHM) programs are highly valuable approaches for equipment reliability. While data-driven methods, particularly Deep Learning (DL), have shown promise, their exclusive reliance on data can lead to false alarms and reduced confidence in the model's prediction by future operators. Hence, Physics-Informed Deep Learning (PIDL) integrates physics-related information into the DL model through an adapted loss function. This study explores the application of PIDL in fault detection for bearings' vibration data, evaluated on data generated by a bearing vibration experimental bench. Data acquisition was performed under 15 Hz rotation speed and includes three operation modes: healthy, light damage and heavy damage. The outlined methodology involves feature preprocessing using envelope analysis and feature selection based on bearing fault characteristic frequencies, DL network construction, and incorporating a damage level threshold model into the network's loss function to guide the training procedure based on expected physical behavior. The proposed framework is also compared to a traditional DL model, offering insights into its effectiveness. The results highlight the significance of combining statistical and physics information for more accurate and physically consistent models in the bearings industry, as the errors produced by the model tend to be more physically consistent and avoid extreme cases that highly damage the model's credibility. This research contributes to developing an approach for constructing and training PIDL models, offering a valuable comparison with traditional DL models and demonstrating the applicability of PIDL in the bearing's domain.

**Keywords:** Physics-Informed Deep Learning, Bearing Vibration, Experimental Vibration Bench, Fault Detection.

## 1. INTRODUCTION

Reliability of equipment and systems is crucial for success and competitiveness in the 21st century's industry. This is achieved through Condition-Based Maintenance policies [1], which manage equipment based on their diagnosed health state, and Prognostics and Health Management programs [2]. Consequently, there is a growing demand for approaches and professionals in reliability engineering. Data storage advances, affordable sensors, and enhanced computer processing power have increased interest in the usability of data-driven methods, particularly Machine Learning and Deep Learning. These methods are well-suited for characterizing complex systems across various fields [3] [4].

ML models, which rely on data to extract patterns and make predictions, have been successfully employed in various tasks, including failure mode classification [5], Remaining Useful Life (RUL) prediction [6] [7], and anomaly detection [8] [9]. However, these models' exclusive dependency on data can lead to false alarms and reduced prediction confidence if the data fails to adequately characterize specific equipment [10]. A major limitation of these models is their lack of integration with physical principles, which can result in misclassifications and reduced credibility. This issue is particularly critical for equipment with rolling elements, such as pumps [4], because bearings are major contributors to system failures, accounting for up to 55% of all failures in rotating machines [10].

In this context, despite numerous contributions regarding classification of rolling elements via Artificial Intelligence (AI) approaches, the aggregation of physical information into those is limited. Thus, there is a growth in interest regarding Physics-Informed Deep Learning methodologies as an attempt to prevent the highlighted limitations. PIDL models, due to physics contribution, tend to minimize predictions and

misclassifications that go against physical principles. Improving this integration between physics and AI can improve maintenance strategies, potentially reducing accidents and economic losses.

Then, this work explores a methodology for Physics-Informed Deep Learning based on [10] for fault detection in bearing vibration data. The approach consists of the steps: 1) feature preprocessing, responsible for structuring the data and extracting physically-meaningful features from the data; 2) DL network construction, where the model is defined; and 3) modified loss function, in which a damage level threshold model is incorporated into the network's loss function to guide the training procedure respecting physical principles.

This study also collected and recorded vibration data on spherical bearings in a mechanical transmission system on an experimental bench, powered by an induction motor operating in three phases. The vibration data was collected and examined using condition monitoring techniques to assess the integrity and performance of the bearings during system operation. The proposed framework is then evaluated on the collected vibration data, while also comparing its effectiveness to a traditional DL model.

The organization of this paper is as follows: section 2 constitutes of a theoretical background regarding the topics related to this work; section 3 introduces the proposed methodology; section 4 presents the results obtained; section 5 concludes the research.

## 2. BACKGROUND

### 2.1. Machine Learning, Neural Networks, and Deep Learning

Machine Learning can be summarized as the creation of algorithms and statistical models enabling computers to enhance their performance on specific tasks through experience. Rather than relying solely on explicit programmer instructions, computers autonomously learn and adapt from data patterns. Mitchell [11] reinforces this by defining machine learning as the improvement in a computer program's performance on a task, as measured by a certain metric, based on the experience (data).

Neural Networks (NN), first named by Warren McCulloch and Walter Pitts in 1943 [12], form a category of models within Machine Learning that emulate the structure and operation of the human brain for data learning purposes. These networks consist of interconnected computational units known as neurons, organized into layers including input, hidden, and output layers. Neurons process input signals using activation functions, producing output signals that propagate through the network. The strength of connections, represented by weights, models the flow of information. During training, Neural Networks adjust their weights by minimizing the disparity between predictions and actual targets in the training data via a process named backpropagation. This iterative process involves optimizing weights using algorithms such as gradient descent to minimize predefined loss functions.

The name Deep Learning represents a subset of the Neural Network field dedicated to constructing and training networks comprising multiple hidden layers. These Deep Neural Networks specialize in executing intricate tasks with remarkable precision by assimilating high-level abstract representations of data, as the hierarchical layer arrangement resulting from increased depth enhances feature extraction capabilities [13]. Diverging from traditional NNs, DL also includes a range of layer types designed to specific objectives. For example, Convolutional layers excel in processing grid-like data, such as images, showing superior performance by extracting spatial hierarchies of features crucial for tasks like image classification. DL's success also results from factors like expansive datasets, enhanced computational capabilities, and sophisticated optimization algorithms. Ongoing research and innovation underscore Deep Learning's significance in addressing real-world challenges under diverse domains.

### 2.2. Physics-Informed Deep Learning

DL methods frequently rely solely on data, neglecting to integrate physical principles alongside statistical techniques, which can lead to subpar model performance, as purely data-centric approaches may not align effectively with the laws of physics. Furthermore, a diminished adherence to physics may correlate with a higher rate of misclassifications that infringe physical principles, consequently reducing the credibility of the learning model as operators might question its accuracy. Physics-Informed Deep Learning methods offer a

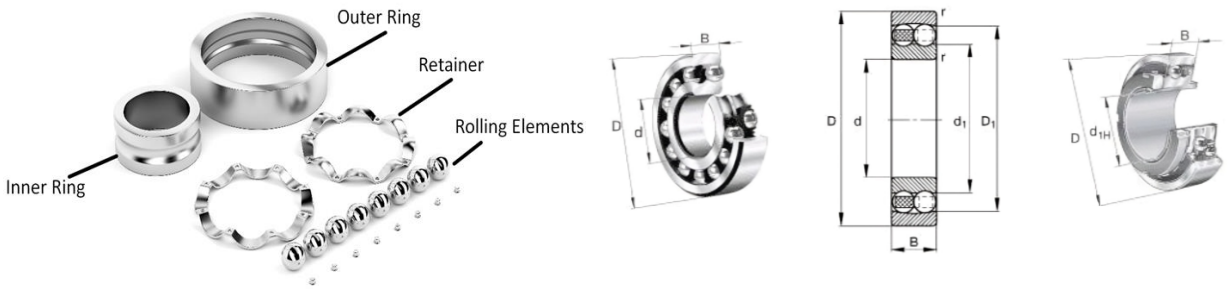
solution to these challenges by addressing misclassifications that violate physics through the incorporation of physical knowledge during its training procedure. The term "Physics-Informed" typically denotes approaches that incorporate physical information by modifying a model's loss function [14].

During the learning process of a Deep Learning model, it iteratively learns through the evaluation of the loss function, which measures the disparity between model predictions and desired outcomes. However, in Physics-Informed training, these loss terms serve a dual purpose. They not only guide the optimization process towards minimizing prediction errors, as in a standard DL model, but also ensure compliance with the underlying physical laws governing the system. This integration of a physical term into the loss function sets it apart by capturing the intrinsic structure of the problem domain, thereby guiding the solution space towards regions that adhere to physical principles. This capability is valuable in scenarios where data may be scarce or noisy, necessitating the explicit inclusion of prior knowledge about the system's behavior for robust learning.

### 2.3. Bearing Vibration Monitoring

Machine vibration is a well-known and reliable way to monitor bearing condition [15]. Bearings are largely responsible for the effective and dependable operation of mechanical transmission systems. Shafts and gears, among other rotating components, can move smoothly because they can support radial and axial loads. Self-aligning bearings are made to handle angular misalignments between the shaft and housing. Because of their special capacity to accept axial and radial misalignments, they are suited for applications where maintaining precise alignment is challenging or impossible. These bearings are often found in machines with flexible shafts or where alignment varies due to vibration, thermal expansion or component wear.

Extensive experiments are needed to gain a deeper understanding of the properties of vibrations in bearings, particularly in failure scenarios, as vibration monitoring has received more attention in recent years and has become more significant [16]. Vibration can be measured with vibration sensors, such as accelerometers and vibration speed transducers [17]. Estimations should be taken on the direction or other structural components that strongly respond to the dynamic force and characterize the overall vibration of the machine. Figure 1 a) depicts a generic bearing consisting of an outside race, inner race, and moving component.




a) Generic Bearing. b) NSK 1205K C3 Bearing.

Figure 1. Rolling Bearing's Components.

The reasons of bearing vibration include variations in external conditions over time between different parts. There are four types of inadequacies that can occur in a moving bearing, depending on where the fault occurs. The purported frequency of bearing defects is determined based on bearing parameters and rotational frequency. Each of these faults is associated with the following formulas used to calculate heading as described in Table 1, where: $N_B$: No. of rolling elements, $B_D$: Ball diameter, $P_D$: Bearing pitch diameter, $\theta$: angle of contact, $RPM$: Rotational speed.

Table 1. Bearing Characteristic Frequencies. Source: [18].

| Frequency name | Description | Equation |
|---|---|---|
| Ball Pass Frequency Outer (BPFO) | Frequency at which failures occur in the outside lane | $RPM * \frac{N_B}{2} * \left(1 - \frac{B_D}{P_D} cos(\theta)\right)$ |
| Ball Pass Frequency Inner (BPFI) | Frequency at which failures occur in the inner lane | $RPM * \frac{N_B}{2} * \left(1 + \frac{B_D}{P_D} cos(\theta)\right)$ |
| Ball Spin Frequency (BSF) | Frequency at which the rolling elements themselves fail | $RPM * \frac{P_D}{B_D} * \left(1 - \left(\frac{B_D}{P_D} cos(\theta)\right)\right)$ |

| | | |
|---|---|---|
| Fundamental Train Frequency (FTF) | Frequency at which a train's cage may fail | $RPM * \frac{1}{2} * \left(1 - \frac{B_D}{P_D} cos(\theta)\right)$ |

According to Randall and Antoni [18], envelope analysis stands as the primary method for bearing diagnosis, given that the raw signal typically lacks informative details about faults. This technique involves a filtering phase to remove frequency bands unrelated to the fault, followed by shifting the signal into the frequency spectrum, which emphasizes the repetitive nature of damage in rotating equipment.

An enhancement to this analysis involves incorporating the Hilbert transform alongside the time-to-frequency transform [10] [19]. This approach proves beneficial as it accentuates local features of the signal, producing an analytical representation of a real-valued signal. In the frequency domain, it introduces a 90º phase shift to all frequency components of a given function, aiding in the detection of instantaneous frequency changes by filtering out rapid oscillations from the signal.

## 3. METHODOLOGY

The considered Physics-Informed DL approach is based on the work of Shen et al. [10] and primarily involves constructing a threshold model based on the expected physical behavior of faults. The output of this supplementary model is then integrated into the DL model's loss function, thereby guiding the training process with crucial physical insights alongside purely statistical considerations. This work follows a standard procedure for AI modeling: 1) feature preprocessing is conducted, encompassing segmentation, transformation, feature selection, train/test partitioning, data labeling, and shuffling; 2) model is constructed, which comprises both i) the DL network and ii) the modified loss function, incorporating the threshold model; 3) classification is executed, and the resultant outcomes are evaluated.

### 3.1. Dataset Description

To collect data, we used a vibration test bench that creates a transmission system and uses an adjustable speed induction motor through a frequency inverter, as seen in Figure 2. Experiments were carried out at the Center for Studies and Tests in Risk and Environmental Modeling (CEERMA), located at the Federal University of Pernambuco (UFPE), Brazil.

Figure 2. Vibration Bench Components.

This study utilized the NSK 1205K C3 bearing type, shown in Figure 1 b). It is classified as a double-row self-aligning deep-groove ball bearing. This sort of bearing, which does not have a seal, is classified as a double-row bearing. Both rows are contained within a spherical raceway located in the outer ring. This design enables the bearing to automatically adjust its alignment. Consequently, any deviation in the alignment of the shaft with respect to the bearing housing does not have a negative impact on performance.

This bearing is well-suited for applications that involve significant misalignment or shaft deflections due to its specific features. Furthermore, self-aligning bearings exhibit the least amount of friction among all types of rolling-element bearings. Even when traveling at higher velocities, the amount of frictional heat generated is little. These bearings are capable of handling high radial loads. They can also be subjected to axial loads in

both directions, although to a lesser degree. The structure of the specific bearing being used is given in Figure 1 b) and its main geometric parameters can be found in Table 2 (left).

Table 2. Specifications for the Bearing (left) and Accelerometer (right).

| **Bearing Specifications** | | **Accelerometer Specifications** | |
|---|---|---|---|
| Inner diameter (d), mm | 25 | Sensitivity | 10.2 mV/(m/s²) |
| Outer diameter (D), mm | 52 | Measurement Range | ± 490 m/s² |
| Pitch diameter (Pd), mm | 38.5 | Frequency Range | 0.5 to 10000 Hz |
| Ball diameter (Bd), mm | 7.14 | Resonant Frequency | 25 Hz |
| No. of rolling elements (Nb) | 12 | Broadband Resolution | 3434 $\mu m$/sec² |
| | | Non-Linearity | ± 1% |
| | | Transverse Sensitivity | ≤ 7% |

The process of measuring data from self-aligning bearings takes place through two accelerometers coupled to the base of the split bearings, which are interconnected by a synchronizer pulley with a guide, after defining the frequency in the inverter, defined as 15 Hz for the experiments in this paper. Vibration signals from the bearings were acquired using the piezoelectric accelerometer, model 603C01, which has the characteristics indicated as shown in Table 2 (right).

The accelerometers were positioned by magnetic method in the split housings where the bearings are located. Due to the methodology of this work, we limited our tests to 3 situations, named healthy, light damage and heavy damage, as shown in Table 3. Figure 3 exemplifies the bearing states for each failure mode.

Table 3. Bearing Damage States.

| **Bearing state** | **Damage extent** |
|---|---|
| Healthy (a) | No damage |
| Light damage (b) | Outer race damage 1mm |
| Heavy damage (c) | Outer race damage 3mm |

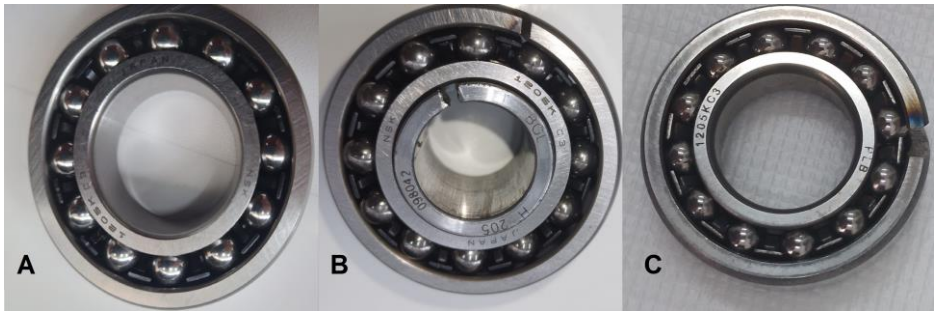


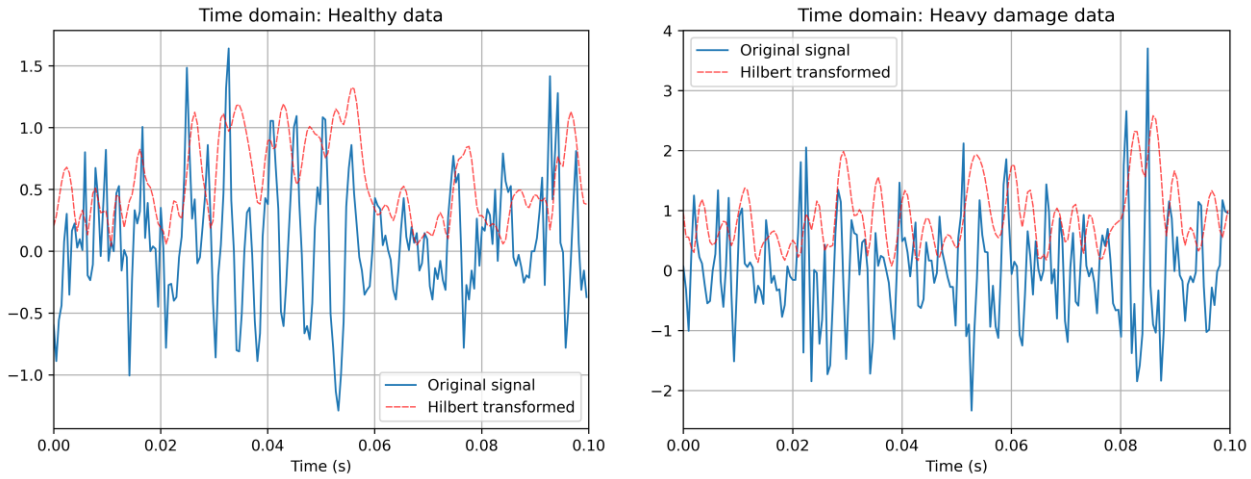

Figure 3. Bearing Damage States: a) Healthy; b) Light Damage; c) Heavy Damage.

The sensor cabling is connected to an amplifier that increases this signal 45x. The sensor measures 0.096mV for every 1g of acceleration. The signal is translated to the input of the Labjack U12 device. Input and output channels are integrated for stimulus-response testing. The sampling rate is defined as 2048 samples/second with a scan rate of 4096 Hz. These values are defined in the LJscope V1.09 software in the number of scans and Scan rate options, with sampling rate set to the maximum possible by this device.

### 3.2. Data Preprocessing

Unlike purely data-driven approaches, Physics-Informed methods integrate known information about the system's expected behavior. In this context, the approach is designed specifically to the system's structure under analysis, making the preprocessing step a crucial component of these models. In this context, our objective is to further investigate the methodology proposed by Shen et al. [10] but adapted to our collected vibration data.

The collected dataset includes multiple time series of vibration data, which requires a data segmentation process, by dividing it into segments of same length [20], in order to feed a DL model, set empirically to 512 points for our dataset. Then, envelope analysis is performed by transforming data from time domain to frequency domain, which is a common procedure when studying bearing damage data, as shown by Lessmeier et al. [21]. We also use the Hilbert transform prior to the domain shift using Fourier transform. The comparison between the Hilbert transformed signal and the original signal in the time domain can be seen in Figure 4.




a) Healthy data b) Heavy damage data

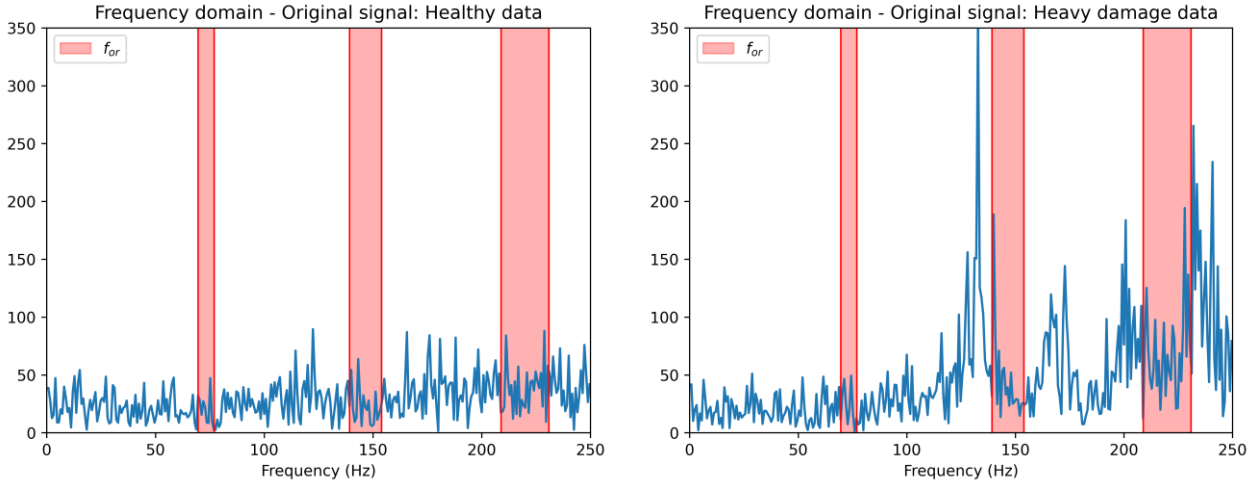
Figure 4. Comparison between the original signal and the Hilbert transformed signal for time domain data.

Based on the formulas shown in Table 1 and the bearing parameters in Table 2, we can calculate the intrinsic fault characteristics for the bearing, considering the operation condition of 15 Hz. Those are given in Table 4.

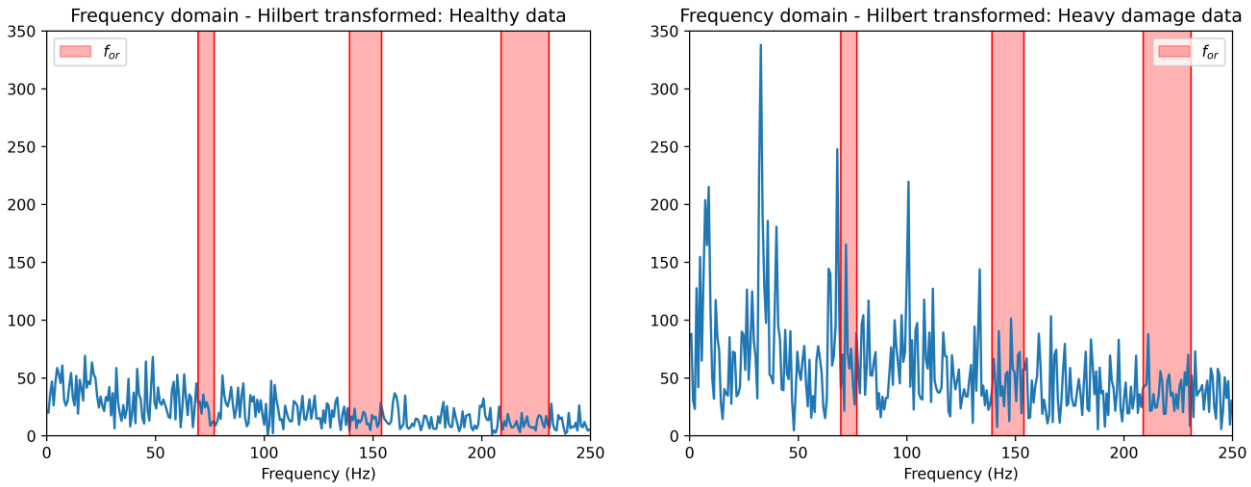
Table 4. Fault Characteristic Frequencies of our bearing.

| **Characteristic Frequency** | **Value (Hz)** |
|---|---|
| Ball Pass Frequency, Outer Race | 73.305 |
| Ball Pass Frequency, Inner Race | 106.695 |
| Ball Spin Frequency | 78.08 |
| Fundamental Train Frequency | 6.10875 |

This work included, alongside each characteristic frequency calculated, up to three harmonics. Also, to better accommodate variations regarding noise and small variations, a sub-band of $\pm 5\%$ is selected for each frequency and harmonic. Data points that are not within any sub-band are discarded. Figure 5 presents examples of frequency domain data, including both the original and Hilbert transformed signals, as well as comparing healthy and heavy damage data. For visualization purposes, only outer race's characteristic frequency and its harmonics are displayed.




a) Original signal, healthy data. b) Original signal, heavy damage data.




c) Hilbert transformed signal, healthy data. d) Hilbert transformed signal, heavy damage data.

Figure 5. Comparison between frequency domain data. The Ball Pass Frequency for the outer race and its harmonics are highlighted in red ($f_{or}$).

This work used a train-test split of 80%-20%, after which the data was shuffled. Number of examples for each subgroup is shown in Table 5.

Table 5. Train and test number of data points and features.

| **Subgroup** | **# Examples** | **# Features** |
|---|---|---|
| $x_{train}$ | 5404 | 32 (points within sub-bands) |
| $x_{test}$ | 1352 | 32 (points within sub-bands) |
| $y_{train}$ | 5404 | 3 (# classes) |
| $y_{test}$ | 1352 | 3 (# classes) |

### 3.3. Network Architecture

This work aims to compare the performance of Physics-Informed Deep Learning model when compared to a traditional Deep Learning model for vibration data collected via our bearing vibration bench. In order to do so, we opt to use a similar model as the one proposed by [10], but we adapted kernel size and stride empirically based on our dataset shape. The final architecture can be visualized in Figure 6.

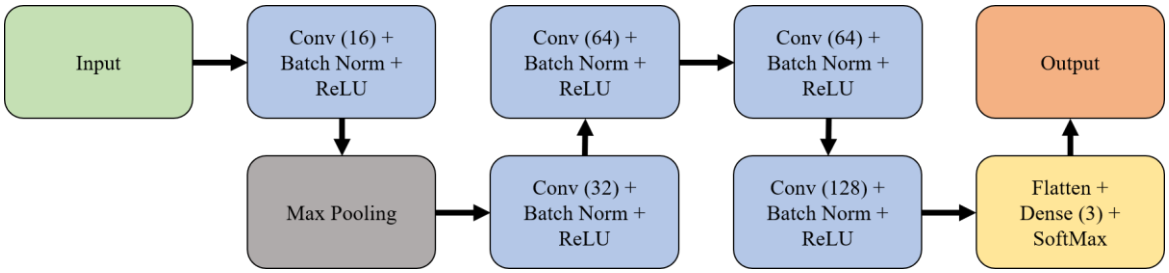



Figure 6. Network Architecture.

### 3.4. Customized Loss Function

Physics-Informed models include the physical dynamics in the model's loss function, as stated in [14]. This addition guides the training process in order to better represent the expected behavior of a system. In this work, following the work of Shen et al. [10], the physical contribution derives from a threshold model, which labels the input based on the highest amplitude among its features, which were selected beforehand by the characteristic frequencies and its harmonics. The threshold model is composed of two thresholds, the first one indicating the maximum amplitude for a data point to be considered healthy, and the second one the maximum amplitude for light damage data, at which point all higher data will be considered heavy damage.

When data points are fed into the CNN model, prediction labels are obtained, which are used to compute the model's loss and subsequently update the parameters of the network. In our case, the threshold model also exists in this process as an additional penalty when the model fails to learn the extreme cases. The procedure for calculating the modified loss is as follows:

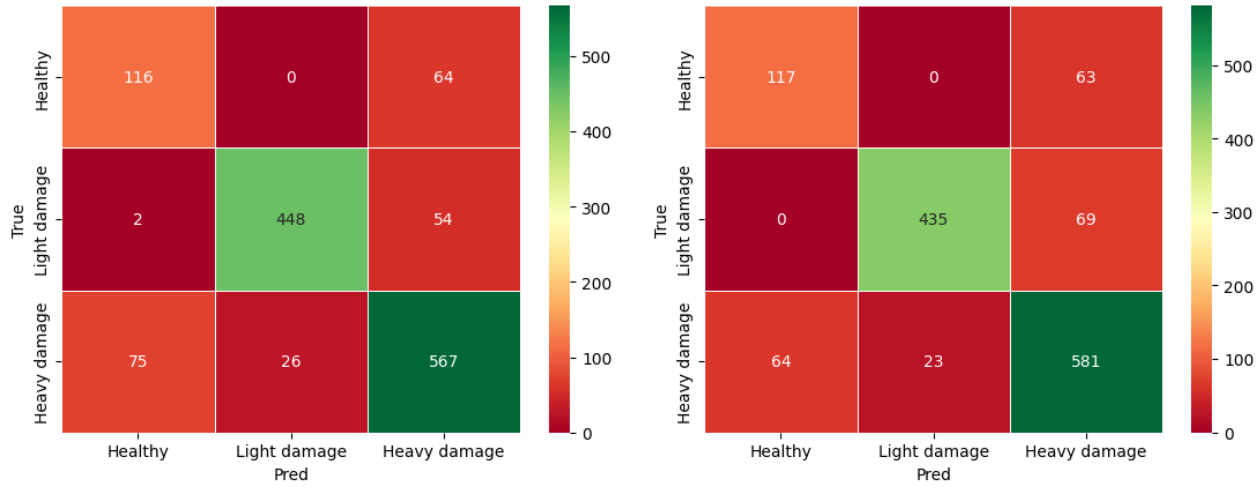
1. The regular CNN loss function, categorical crossentropy (CCE), is calculated for each example, based on the label predicted by the model $y_{pred}$ and the true known label of the same input $y_{true}$;
2. Given the amplitude of the input, it is labeled by the threshold model $y_{th}$;
3. A second CCE loss function is calculated for each example, now based on the label predicted by the original CNN model *y_pred* but using the threshold output $y_{th}$ as the 'true label';
4. The threshold model's CCE is filtered:
   4.1. Entries where $y_{th}$ predicts healthy, but $y_{pred}$ does not predict healthy are multiplied by $\alpha$;
   4.2. Entries where $y_{th}$ predicts heavy damage, but $y_{pred}$ does not predict as such are multiplied by $\beta$;
   4.3. All other entries are multiplied by 0 (no penalty added);
5. Both CCE's, the traditional one from the regular CNN model and the one obtained by the threshold model, are added into the final loss function.

For our dataset, the first and the second threshold values were identified as 53.6345 and 110.8469, respectively. The values for $\alpha$ and $\beta$, responsible for tuning the penalty amount based on physical expected behavior, were set to 0.05 (5% penalty). As already stated, a regular model – with only the regular CCE loss function – is also implemented in order to compare the results. Both models were trained for 50 epochs using a batch size of 64, ADAM optimizer and initial learning rate of 0.001.

## 4. RESULTS

This work implemented a Physics-Informed Deep Learning model for fault detection in bearing vibration data from a vibration bench. This way, the desired output is a label, representing the class in which the input data comes from. Also, the same model is implemented but without using the physics-informed loss term, in order to compare the results obtained and better comprehend the gains related to the physical addition.

The confusion matrices for both the regular and the Physics-Informed DL trained models is shown in Figure 7. Although the general accuracy gains between both models did not improve much – 83.6% for the regular model against 83.8% for the physics-informed one –, it is visible to note the physical guidance's impact related to the extreme cases: all the data from true labels healthy and heavy damage got improved, in regards to light damage label, especially heavy damage being classified as healthy, dropping from 75 to 64 occurrences, or 14.7% reduction. This is as such due to the threshold model additionally penalizing errors in those classes specifically, as they tend to have a much higher impact in terms of model credibility and false alarms than errors related to light damage. In terms of light damage inputs, the physics informed classified more data as heavy damage compared to the traditional CNN model, increasing misclassifications from 54 to 69, a 27.8% increase.




a) Regular DL model: 83.6% accuracy. b) PIDL model: 83.8% accuracy.

Figure 7. Confusion Matrices of Trained Models.

Even though the increase in misclassifications related to light-heavy damages is higher than the heavy-healthy reduction, it has a much lesser impact in the model credibility, as operators warned by the model will actually find damages in the bearing. In the case of heavy damage classified as healthy, operators would not know about the damage existing until the machine breaks, resulting in high cost and maybe even risk of accidents. Thus, the additional loss term, designed specifically to reduce the most extreme misclassifications, impacted substantially in the model's training guidance.

The models were trained in a simple computer, containing Intel® Core™ i5-5200U CPU @ 2.20GHz, 8GB RAM, NVIDIA GeForce 920M, with the training procedure taking around 1m30s for the regular model and 1m45s for the Physics-Informed Deep Learning model. Thus, its computation is efficient, which may vary based on dataset sizes, as our tests regarding the vibration bench just started, so our dataset size is relatively small.

## 5. CONCLUSION

Artificial Intelligence and Deep Learning approaches highly benefit Condition-Based Maintenance and Prognostics and Health Management approaches. Still, those models are purely reliant on data, which could possibly drive a reduction in the confidence of the models by the operators, when false alarms occur in contradiction with the expected physical behavior.

A Physics-Informed Deep Learning was implemented and evaluated, including physical knowledge regarding a system into the model's loss function, alongside the statistical information from the data. For our case, the physics-informed term comes in the form of a threshold model, responsible for predicting, based on the input's frequency-domain amplitude, a damage label for the data. This information is then used to guide the extreme cases – healthy and heavy damage misclassifications – in order to address the limitations stated.

Results show a slightly increased accuracy when comparing the traditional and the physics-informed DL models: 83.6% and 83.8%, respectively. The PIDL also show a decrease in all misclassifications related to healthy and heavy damage true labels, which are the main contributors to a reduction in the model's confidence. On the other hand, light damage into heavy damage misclassifications increased, but those impact the model's credibility in a much lesser degree, as operators that inspect the machine will actually find damage in the bearings. So, there is a potential reduction in the number of false alarms when using Physics-Informed Deep Learning models. Thus, the methodology has potential to be further evaluated in other scenarios.

This paper contributes mainly as i) implementing a procedure for training deep learning models guided by physical knowledge of bearings; ii) comparing results of pure-statistical models with statistical-physics hybrid models; iii) example of application of PIDL in data generated using an experimental bearing vibration bench, which simulates rotating machinery widely available in the industry.

Some future steps are highlighted: i) gather knowledge regarding our newly acquired vibration bench in order to gather more data under different operation conditions, damage levels, bearing types and fault modes; ii) further study physics-informed approaches, specially in the context of rotating machinery, in order to develop different approaches for custom loss functions; iii) consider different preprocessing techniques, such as STFT, mel scale for acoustic data, and others.

**Acknowledgements**

The authors thank the following Brazilian research funding agencies for the financial support: Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE), APQ-1101-3.08/21, APQ-0394-3.08/22; Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), 308924/2021-5, 310892/2022-8, 409701/2022-0; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Finance Code 001. We also thank the Human Resources Program 38.1 (PRH 38.1) entitled Risk Analysis and Environmental Modeling in the Exploration, Development and Production of Oil and Gas, financed by Agência Nacional de Petróleo, Gás Natural e Biocombustíveis (ANP) and managed by Financiadora de Estudos e Projetos (FINEP), 044819.

## References

[1] JARDINE, A. K. S.; LIN, D.; BANJEVIC, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. **Mechanical Systems and Signal Processing**, 20, I. 7, pp. 1483-1510. DOI: https://doi.org/10.1016/j.ymssp.2005.09.012.

[2] PECHT, M. G. (2009). Prognostics and Health Management of Electronics. **Encyclopedia of Structural Health Monitoring**, pp. 1-13. DOI: https://doi.org/10.1002/9780470061626.shm118.

[3] SOOMRO, A. A. et al. (2022). Integrity assessment of corroded oil and gas pipelines using machine learning: A systematic review. **Engineering Failure Analysis**, 131, Article nº 105810. DOI: https://doi.org/10.1016/j.engfailanal.2021.105810.

[4] ALIYU, R.; MOKHTAR, A. A.; HUSSIN, H. (2022). Prognostic Health Management of Pumps Using Artificial Intelligence in the Oil and Gas Sector: A Review. **Applied Sciences - Basel**, 12, I. 22, Article nº 11691. DOI: https://10.3390/app122211691.

[5] YUAN, L. et al. (2020). Rolling Bearing Fault Diagnosis Based on Convolutional Neural Network and Support Vector Machine. **IEEE Access**, 8, pp. 137395-137406. DOI: 10.1109/ACCESS.2020.3012053.

[6] MAIOR, C. B. S. et al. (2016). Remaining Useful Life Estimation by Empirical Mode Decomposition and Support Vector Machine. **IEEE Lain America Transactions, v. 14, no. 11**, pp. 4603-4610.

[7] SIKORSKA, J. Z.; HODKIEWICZ, M.; MA, L. (2011). Prognostic modelling options for remaining useful life estimation by industry. **Mechanical Systems and Signal Processing**, 25, I. 5, pp. 1803-1836. DOI: https://doi.org/10.1016/j.ymssp.2010.11.018.

[8] BAKDI, A.; KOUADRI, A. (2017). A new adaptive PCA based thresholding scheme for fault detection in complex systems. **Chemometrics and Intelligent Laboratory Systems**, 162, pp. 83-93. DOI: https://doi.org/10.1016/j.chemolab.2017.01.013.

[9] ZHOU, C.; PAFFENROTH, R. C. (2017). Anomaly Detection with Robust Deep Autoencoders. **Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining**, pp. 665–674. DOI: https://doi.org/10.1145/3097983.3098052.

[10] SHEN, S. et al. (2021). A physics-informed deep learning approach for bearing fault detection. **Engineering Applications of Artificial Intelligence**, 103, Article nº 104295. DOI: https://doi.org/10.1016/j.engappai.2021.104295.

[11] MITCHELL, T. M. (1997). Machine Learning. **McGraw Hill, New York**, pp. 421.

[12] MCCULLOCH, W. S.; PITTS, W. (1943). A Logical Calculus of the Ideas Immanent in Nervous Activity. **Bulletin of Mathematical Biophysics**, 5, pp. 115-133. DOI: https://doi.org/10.1007/BF02478259.

[13] GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. (2016). Deep Learning. **MIT Press**. Available in: <http://www.deeplearningbook.org>.

[14] THUEREY, N. et al. (2021). Physics-based Deep Learning. **WWW**. Available in: <https://physicsbaseddeeplearning.org>.

[15] JAYASWAL, P.; WADHWANI, A.; MULCHANDANI, K. (2008). Machine Fault Signature Analysis. **International Journal of Rotating Machinery**, Article nº 583982. DOI: https://doi.org/10.1155/2008/583982.

[16] ABBASION, S. et al. (2007). Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine. **Mechanical Systems and Signal Processing**, 21, I. 7, pp. 2933-2945. DOI: https://doi.org/10.1016/j.ymssp.2007.02.003.

[17] SAFIZADEH, M. S.; LATIFI, S. K. (2014). Using multi-sensor data fusion for vibration fault diagnosis of rolling element bearings by accelerometer and load cell. **Information Fusion**, 18, pp. 1-8. DOI: https://doi.org/10.1016/j.inffus.2013.10.002.

[18] RANDALL, R. B.; ANTONI, J. (2011). Rolling element bearing diagnostics — A tutorial. **Mechanical Systems and Signal Processing**, 25, I. 2, pp. 485-520. DOI: https://doi.org/10.1016/j.ymssp.2010.07.017.

[19] KANARACHOS, S. et al. (2017). Detecting anomalies in time series data via a deep learning algorithm combining wavelets, neural networks and Hilbert transform. **Expert Systems with Applications**, 85, pp. 292-304. DOI: https://doi.org/10.1016/j.eswa.2017.04.028.

[20] ZHANG, S. et al. (2019). Semi-Supervised Learning of Bearing Anomaly Detection via Deep Variational Autoencoders. **ArXiv eprint 1912.01096**. DOI: https://doi.org/10.48550/arXiv.1912.01096.

[21] LESSMEIER, C. et al. (2016). Condition Monitoring of Bearing Damage in Electromechanical Drive Systems by Using Motor Current Signals of Electric Motors: A Benchmark Data Set for Data-Driven Classification. **Proceedings of the European Conference of the PHM Society**, 3, Article nº 1. DOI: https://doi.org/10.36001/phme.2016.v3i1.1577.