Novelty Detection in Rotating Machinery: Assessment of Unsupervised Machine Learning Models for Medium-Sized Industrial Bearings

Luigi Gianpio Di Maggio
Dept. of Mechanical and Aerospace Engineering
Politecnico di Torino
Torino, Italy
luigi.dimaggio@polito.it

Eugenio Brusa
Dept. of Mechanical and Aerospace Engineering
Politecnico di Torino
Torino, Italy
eugenio.brusa@polito.it

Cristiana Delprete
Dept. of Mechanical and Aerospace Engineering
Politecnico di Torino
Torino, Italy
cristiana.delprete@polito.it

***Abstract*—Early anomaly detection in industrial rotating equipment is crucial for enhancing predictive maintenance methodologies and preventing unexpected failures. However, real-world industrial settings often lack structured fault-labeled datasets, making supervised classification approaches impractical. In such contexts, unsupervised novelty detection methodologies emerge as the most viable alternative, as they do not require labeled fault data and operate under the assumption that only normal conditions are available during training.**

This research investigates the application of novelty detection techniques for industrial bearing condition monitoring, comparing machine learning-based methods with fixed threshold-based strategies derived from standards. The analysis is conducted on a dedicated experimental dataset, representing the first contribution of its kind for spherical bearings with localized damage, tested under variable speed and load conditions.

The findings reveal that fixed-threshold strategies result in an excessive number of false positives, limiting their practical applicability in industrial monitoring systems. Among the assessed machine learning algorithms, Isolation Forest achieved the highest recall, detecting the largest number of anomalies, while Local Outlier Factor (LOF) demonstrated superior precision, accuracy, F1-Score and precision-recall AUC.

***Index Terms*—novelty detection, rotating machinery, bearings, condition monitoring, vibration analysis, unsupervised learning**

I. Introduction

Industrial rotating machinery can greatly benefit from the adoption of predictive maintenance technologies, which optimize resource utilization and generate significant economic returns by reducing unplanned downtime and preventing catastrophic failures. These advantages justify the investment in acquisition, processing, and analysis of vibration signals for condition monitoring and early fault detection [1]–[3].

Over the past decade, there has been an exponential increase in research focusing on the use of machine learning (ML) and deep learning (DL) techniques for vibration signal analysis, particularly using accelerometric data, to diagnose bearing defects and predict maintenance needs in advance [4], [5]. However, a significant portion of existing works is centered around fault diagnosis algorithms [6]–[8], which typically require a substantial amount of labeled data covering different damage conditions and operational states. In real-world industrial settings, structured datasets with labeled faults are rarely available, limiting the practical applicability of such approaches.

Some studies have explored anomaly detection methodologies, which aim to identify deviations from normal operational behavior in datasets that may not contain labeled fault data. De Fabritiis and Gryllias [9] leveraged unlabeled vibration data to learn a distance metric and effectively discriminate anomalous samples from normal ones in degradation stages, Sangeetha et al. [10] adopted hybrid SVM-XGBoost model optimized by Henry Gas Solubility Optimization for anomaly detection in rotary equipment, achieving 94% accuracy, Mishra et al. [11] employed Motor Current Signature Analysis (MCSA) for system-level multi-fault anomaly detection using machine learning algorithms and Asavalertpalakorn et al. [12] investigated novelty detection method was implemented as a one-class classification to detect early signs of failure in a rolling bearing using long short-term memory (LSTM) autoencoder. However, most of these studies rely on benchmark datasets [13], [14] that are collected under highly controlled conditions, often focusing on small-scale bearings, and do not account for medium-to-large-sized bearings operating under variable speed and load conditions, which are common in heavy industrial environments.

This study investigates the applicability of novelty detection techniques for industrial systems, under the assumption that no labeled fault data is available for training, thereby necessitating the use of unsupervised learning approaches. To achieve this, a comparative analysis is conducted between state-of-theart novelty detection algorithms and fixed-threshold vibration assessment methods, such as those derived from ISO 20816 standards [15]–[17]. The goal is to evaluate adaptive machine learning models and evaluate their performance with respect to non-adaptive threshold-based approaches in detecting anomalies in real-world scenarios.

Additionally, this work presents a dedicated experimental campaign and an exclusive dataset developed at the Politecnico di Torino [18], specifically designed for the study of medium-to-large-sized damaged bearings. The dataset [19] includes vibration signals recorded under variable axial and radial loads and across multiple rotational speeds, providing a unique resource for evaluating novelty detection techniques in realistic industrial conditions.

II. Novelty Detection in Industrial Equipment

This section presents the methodology and dataset employed for the assessment of novelty detection techniques within the context of industrial bearing condition monitoring. Initially, the experimental configuration and the data acquisition methodology are introduced, emphasizing the specifications of the test rig and the vibrational dataset utilized. Subsequently, the foundational concepts of unsupervised novelty detection are articulated as well as the selection of machine learning algorithms alongside the heuristic ISO threshold-based approach.

A. Vibration Dataset for Medium-Sized Industrial Bearings

The comparative evaluation proposed in this paper was carried out by means of a database of vibrational signals built for the purpose. Data were extracted from the experimental campaign pertinent to the dataset [19] obtained through the medium-sized bearing testing apparatus located within the mechanical engineering laboratories of the Politecnico di Torino. A comprehensive description of the testing apparatus is presented in the works [18], [20], [21], while the arrangement of the dataset is elucidated in [19].

The test rig shown in Fig. 1 and Fig. 2 is characterized by a modular architecture that facilitates the use of bearings with outer diameters variable up to 420 mm, rendering it a multifaceted instrument for the assessment of medium- and large-diameter bearings' performance. The test rig is equipped with an inverter-controlled 30 kW electric motor SIEMENS®, which allows variable rotational velocities. Furthermore, the system encompasses independent hydraulic actuators for applying radial and axial forces up to 200 kN.

Vibrational data are captured using high-sensitivity piezoelectric accelerometers SKF®CMSS 2200T affixed to rigid bearing adapters. Additionally, the entire system is overseen by specialized TestLab software that facilitates real-time management of test variables and supervision of bearing operational conditions.

In the current investigation, the SKF®22240 CCK/W33 bearing (Fig. 3) was subjected to four distinct operational scenarios: no damage (H), damage to the inner raceway (IR), damage to the outer raceway (OR), and damage to the rolling element (B). Damage was artificially induced via chip removal,

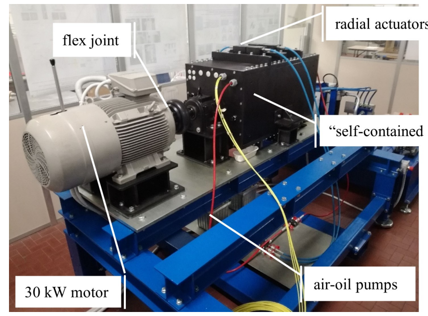


Fig. 1: Test rig for bearing monitoring: self-contained box [18].

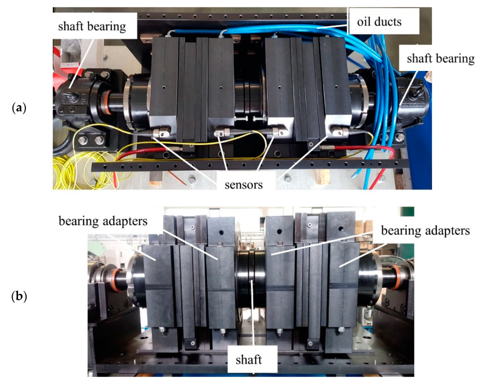


Fig. 2: Test rig for bearing monitoring: shaft and adapters [18].

with a diameter of 2 mm and a depth of 0.5 mm. These dimensions are comparable in scale with those present in the literature dataset.Test conditions are reported in Table I.

TABLE I: Load Cases and Nominal Speeds.

	Case 1	Case 2	Case 3	Case 4
Radial load (kN)	0	62.4	124.8	124.8
Axial load (kN)	0	0	0	49
Nominal speeds (rpm)	127, 227, 353, 457, 523, 607, 727, 877, 937, 997			

For each experiment, vibrational signals of 30 seconds duration were recorded at a sampling frequency of 20480 Hz. The signals were divided into segments comprising 4096 samples, with a 50% overlap between successive segments. Subsequently, the dataset was established by extracting the following features in the temporal domain: Mean T1, Standard Deviation T2, Root Amplitude T3, RMS T4, Absolute Maximum T5, Skewness T6, Kurtosis T7, Crest Factor T8, Form Factor T9, Shape Factor T10, and Impulse Factor T11.

To the authors' knowledge, the testing apparatus and the dataset constitute the first experimental contribution capable of furnishing vibrational data for medium and large bearings, inclusive of axial loading conditions. In comparison to well-known datasets previously documented in the literature [13], [14], which predominantly concentrate on small bearings and circumscribed operational conditions, the present study delineates a distinctive dataset of industrial significance for applications within the heavy industry sector.

(a) Disassembled bearing. (b) IR damage [21].

Fig. 3: Bearing SKF®22240 CCK/W33.

B. Unsupervised Novelty Detection

The current study explores the anomaly detection efficacy of a range of predictive algorithms grounded in non-deep machine learning techniques. In this initial examination, deep learning algorithms and neural networks are deliberately excluded, as the authors posit that such approaches would introduce an extraneous level of complexity (e.g., the selection of neuron counts, layer configurations, and other architectural considerations) that could obfuscate the interpretation of the findings and introduce additional dependent variables into the comparative analysis.

The investigation is undertaken with the understanding that in real industrial settings, labeled datasets are frequently unavailable for the execution of supervised classification. Consequently, it is often pragmatic to employ unsupervised techniques that do not necessitate an extensive labeling by the operators. This consideration is particularly salient in industrial applications where the presence of abnormal data is constrained and often fails to represent the spectrum of potential anomalies. A further pivotal component is the differentiation between novelty detection and outlier detection, which is directly mirrored in the methodology employed in this study.

Novelty Detection is concerned with the identification of new instances that were not included during the training phase and that diverge from the established normal training data [22]. This methodology develops on a training dataset devoid of anomalies, rendering it especially appropriate for industrial contexts characterized by the rarity and unpredictability of anomalies. Conversely, Outlier Detection concerns the identification of data points that significantly deviate from the predominant data within the training dataset [23], which may already be included in the training phase. In our analysis, we situate ourselves within the framework of novelty detection, wherein the training dataset remains uncontaminated by outliers. Thus, the objective is to assess the efficacy of the algorithms in recognizing novelties that manifest within the test dataset.

The dataset was partitioned into a training subset, which exclusively comprises normal data, and a testing subset, which encompasses both normal data and anomalies. The methodology employed for this partitioning is delineated below.

- Training subset: comprises solely normal data sourced from the original dataset.
- Testing subset: encompasses both normal data and anomalies, with the proportion of anomalies established at 10% of the total testing subset (Fig. 4).

Table II shows the dataset partitioning and evaluation setup. Namely, the initial dataset had 47,840 samples, the training subset size (normal data exclusively) is of 5,980 samples, the testing subset size (normal + anomalous) is of 5,980 samples, where 5,382 are normal data and 598 are anomalies. Anomalies are randomly drawn and can therefore belong to any damage class. The tested algorithms underwent evaluation across five repetitions to enhance statistical significance.

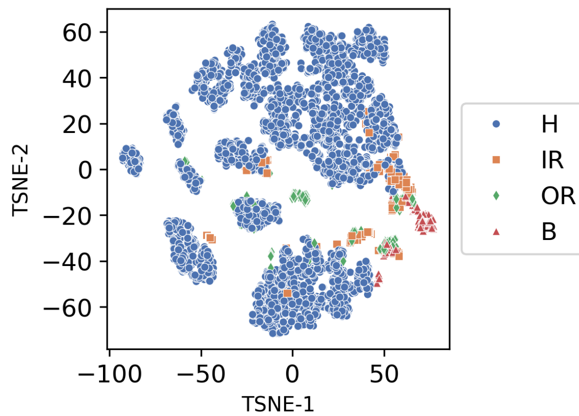


Fig. 4: tSNE representation of test data on fifth repetition.

TABLE II: Dataset Description and Evaluation Setup.

Description	Samples
Initial dataset	47,840
Training subset (normal data)	5,980
Testing subset size	5,980
- Normal data	5,382
- Anomalies	598
Evaluation repetitions	5

For the purpose of this analysis, the following algorithms were investigated: Isolation Forest [24], One-Class SVM [25], Local Outlier Factor (LOF) [26], and Elliptic Envelope [27]. These algorithms are chosen as they are among the most common for unsupervised novelty detection and do not require anomalies in the training set.

- Isolation Forest: This algorithm is particularly suitable for anomaly detection as it isolates observations by randomly selecting features and splitting values. Its efficiency and scalability make it effective for high-dimensional datasets, enabling it to perform well in novelty detection scenarios.
- One-Class SVM: This method constructs a decision boundary around the data distribution, aiming to identify novel or anomalous samples that fall outside the boundary. Its kernel-based approach allows it to model complex data distributions, making it versatile for novelty

detection. In this study, radial basis function (RBF) kernel with auto γ and $\nu = 0.1$ is employed.

- Local Outlier Factor (LOF): LOF evaluates the local density deviation of a sample with respect to its neighbors. Samples with significantly lower densities are flagged as anomalies, making it effective for identifying local patterns and deviations. LOF is set in novelty detection mode with a number of neighbors of 20.
- Elliptic Envelope: This algorithm assumes the data follows a Gaussian distribution and identifies anomalies based on the Mahalanobis distance. Although more tailored for traditional outlier detection tasks, it was incorporated here for comparative purposes. Elliptic Envelope is set with contamination factor 0.1.

In addition to the aforementioned algorithms, a heuristic methodology grounded in the ISO 20816-1 [15] and 20816-3 [16] standards was also evaluated. This methodology leverages the energy of the vibration signal, as indicated by the root mean square (RMS) broadband velocity, to detect anomalies in relation to a predetermined threshold. The application of this standard carried out by the authors in this work is fundamentally heuristic and does not rigorously adhere to procedures and considerations delineated by the standard itself but rather utilizes the threshold concept to assess the efficacy of methods predicated on acceleration signal energy. Namely, the RMS vibration velocities values were computed to assess the operational condition of the machines being investigated. A benchmark of 4.5 mm/s, as indicated by the ISO 20816-1 and ISO 20816-3 standards [15], [16], was employed to activate condition monitoring alerts. This benchmark could generally delineate the parameters of acceptable operational conditions from those necessitating intervention in industrial machinery. Fig. 5 depicts the probability distributions of RMS vibration velocity corresponding to various health states of the bearings being analyzed.

Frequency-based methods, relying on characteristic fault frequencies are also widely used in bearing fault diagnosis. Since this paper focuses on unsupervised novelty detection, the authors did not consider prior knowledge of fault types or spectral signatures. This agnostic approach is particularly beneficial in industrial settings where anomalies may arise without predefined fault models or in unconventional frequency patterns. While we acknowledge the effectiveness of frequency-based methods, our goal is to develop more versatile novelty detection techniques applicable to dynamic and variable real-world scenarios.

III. Results and Discussion

This study aims to evaluate the effectiveness of of the above-mentioned novelty detection algorithms for condition monitoring in industrial bearings. The primary objective is to compare the performance of machine learning-based approaches against a threshold-based method derived from standards. The evaluation is based on key classification metrics, including accuracy, precision, recall, F1-score, and Precision-Recall Area Under the Curve (PR AUC) to determine the most reliable approach for detecting bearing anomalies under varying operating conditions. PR AUC is particularly relevant given the highly imbalanced nature of the dataset, as it provides a more informative assessment of model performance in scenarios with a significantly lower prevalence of anomalies.

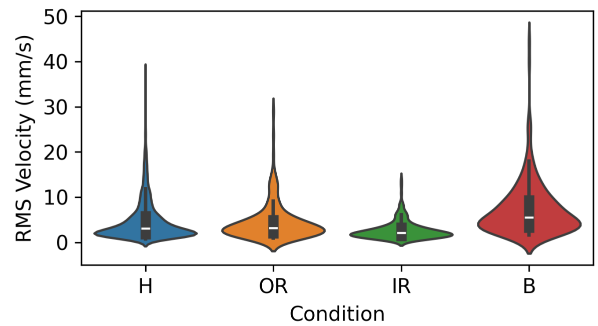


Fig. 5: Distribution of RMS velocities on test data.

A primary significant finding pertains to the methodology based on the ISO 20816 threshold, which utilizes a threshold of 4.5 mm/s to differentiate between normal and anomalous states. Despite this approach exhibiting a sensitivity (Fig. 7c) that is similar with that of poor performing models such as the Elliptic Envelope (Fig. 6a), it manifests an exceedingly elevated incidence of false positives, as delineated by the confusion matrix shown in Fig. 6b and overall recall reported inFig. 7c. This phenomenon results in a markedly diminished precision, rendering an uncalibrated threshold suboptimal within the framework of an automated industrial monitoring system, wherein minimizing false alarms is paramount to avert unnecessary interruptions and maintenance activities.

It becomes evident that the capacity to accurately identify damages when they are indeed present (i.e., sensitivity or recall) is superior in the Isolation Forest (Fig. 6c and Fig. 7c), which succeeded in detecting the greatest number of real anomalies. Nonetheless, this achievement is realized at the cost of a considerable number of false positives, thereby adversely impacting precision (Fig. 7b). To mitigate the high false positive rate observed in Isolation Forest, several strategies could be considered, including fine-tuning the decision threshold to optimize the trade-off between recall and precision, adjusting hyperparameters such as the number of estimators and sample size, leveraging ensemble methods to enhance model robustness, and incorporating additional features to improve the discrimination between normal and anomalous conditions. Conversely, the Local Outlier Factor (LOF) (Fig. 6d) algorithm emerges as the one that recorded the highest precision (Fig. 7b), indicating that when the model predicts an anomaly, it is correlated with the highest likelihood of actual presence. This is congruent with the observation that LOF possesses the least number of false positives, thereby representing the most conservative approach in declaring an anomaly. Furthermore, LOF attained the highest overall accuracy value Fig. 7a, which quantifies the aggregate proportion of correct predictions relative to the total sample size.

Another critical metric for assessing the effectiveness of a model in a novelty detection context is the F1-score shown in Fig. 7d, which encapsulates the equilibrium between precision and recall. The LOF model achieved the optimal F1-score, signifying that it is the model that most effectively reconciles the capability to identify anomalies while mitigating false positives. A high F1-score suggests that the model is proficient in minimizing overall errors while sustaining both remarkable sensitivity and accuracy.

Furthermore, PR AUC, shown in Fig. 7e, further supports the effectiveness of the LOF model. LOF achieves the highest PR AUC, reinforcing its ability to maintain a strong balance between precision and recall. Given the dataset's strong class imbalance, PR AUC provides a more insightful evaluation of the model's capacity to detect anomalies without being disproportionately influenced by the prevalence of normal samples. This result confirms that LOF not only minimizes false positives but also consistently identifies true anomalies, making it the most robust approach in this context.

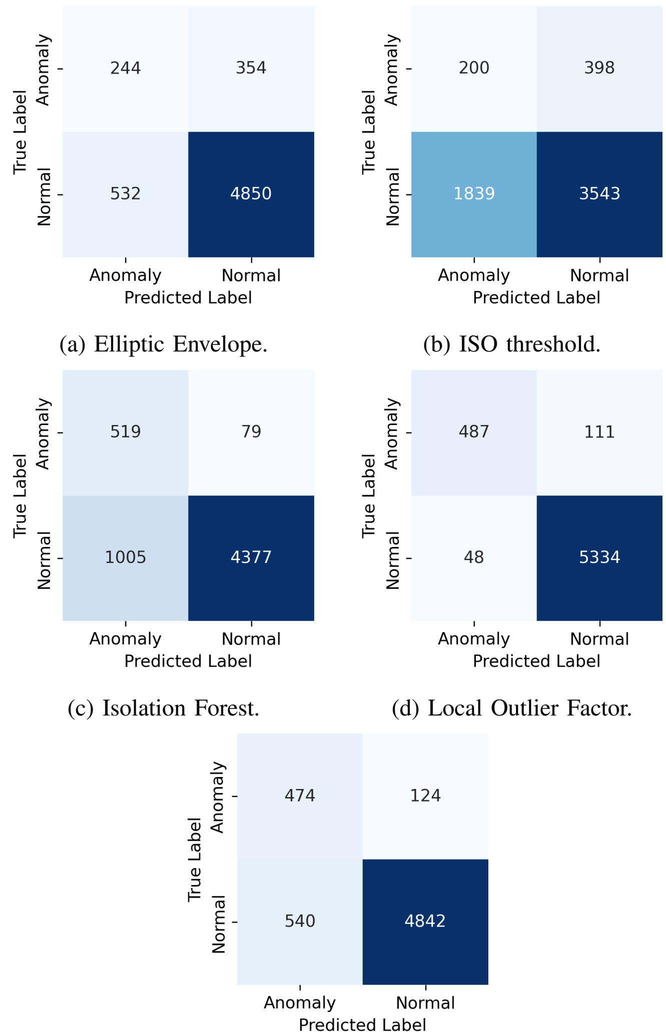
The findings, derived from five independent repetitions, reveal a low variability in the performance of the models, suggesting a degree of robustness in the advanced evaluations. Nevertheless, to enhance the level of statistical significance, further studies carried out on more extensive samples and additional random iterations will be required, in order to more comprehensively assess the variability of the predictions. An additional factor to contemplate is that the anomalies were randomly selected from signals exhibiting diverse operational conditions, spanning various categories of damage along with differing speed and load regimes. This variability could potentially influence the performance of the algorithms, rendering certain models more responsive to specific operational conditions. Consequently, a subsequent phase will involve the exploration of adaptive strategies that facilitate the customization of detection based on the distinct characteristics inherent to each operational regime.

In summary, the results confirm that fixed thresholds based on ISO standards may be inadequate for advanced condition monitoring applications, while machine learning-based methods, especially LOF and Isolation Forest, offer a better compromise between accuracy and ability to detect real anomalies.

IV. Conclusions

This research investigated novelty detection algorithms pertinent to the monitoring of industrial bearing conditions, juxtaposing machine learning-based methodologies with a threshold-based approach extrapolated from ISO 20816 standards. The findings underscore significant differences in performance metrics, accentuating the inadequacy inherent in static threshold techniques while exposing the advantages of adaptive machine learning strategies.

Furthermore, this study presents an experimental campaign that, for the first time, explores novelty detection applied to medium-sized spherical bearings with localized damage, operating under variable loads and rotational speeds. This aspect is particularly relevant, as existing studies often focus on small-scale bearings or controlled conditions, whereas the present work addresses a more complex and industrially significant scenario.

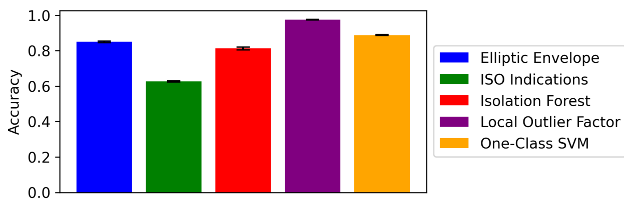


(a) Elliptic Envelope. (b) ISO threshold. (c) Isolation Forest. (d) Local Outlier Factor. (e) One-Class SVM.

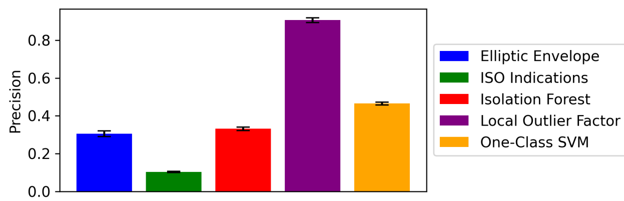
Fig. 6: Confusion matrices.

The results demonstrate that general thresholding could produce a substantial false positive rate, markedly diminishing precision. Although this approach attains a sensitivity that is commensurate with simplistic statistical models such as the Elliptic Envelope, it is deficient in robustness within practical scenarios where the minimization of false alarms is of paramount importance.

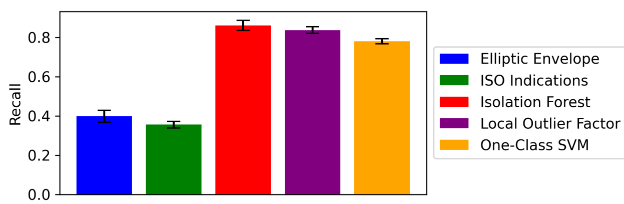
Among the evaluated machine learning algorithms, the Isolation Forest algorithm manifested the highest sensitivity, proficiently identifying a considerable quantity of anomalies. Nevertheless, this elevated sensitivity was accompanied by a substantial incidence of false positives, thereby impairing its overall precision. In contrast, the Local Outlier Factor (LOF) algorithm exhibited the highest precision and comprehensive



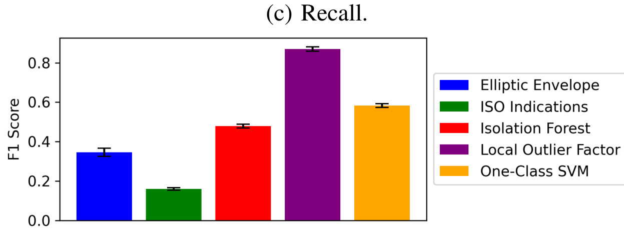
(a) Accuracy.



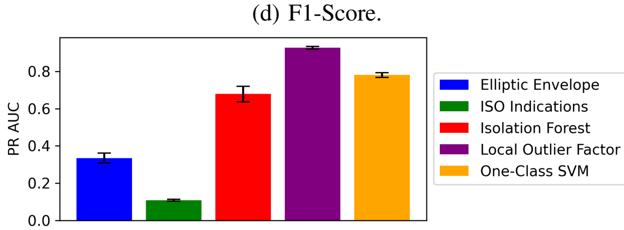
(b) Precision.



(c) Recall.



(d) F1-Score.



(e) PR AUC.

Fig. 7: Overall results.

accuracy, rendering it the most conservative model in predicting anomalies and curtailing false alarms.

Moreover, the LOF attained the highest F1-score and PR AUC, indicative of the optimal equilibrium between precision and recall. This observation implies that while LOF is recognized as the most accurate model and balanced in terms of precision and recall, whereas the Isolation Forest may represent the most sensitive to anomalies. The results, derived from five independent iterations, reveal a low variance, thereby reinforcing the statistical reliability of the conclusions drawn.

Notwithstanding the robustness of the models assessed, further validation utilizing larger datasets with an increased number of random iterations should be performed to assess statistical significance. Additionally, adaptive detection strategies customized to particular operational conditions could further enhance model performance by mitigating sensitivity to fluctuations in speed and load. Lastly, an ensemble methodology, which amalgamates multiple models, could serve to optimize both precision and recall, thereby achieving a more dependable novelty detection system for industrial applications.

References

[1] A. R. Mohanty, *Machinery condition monitoring: Principles and practices*. CRC Press, 2014.

[2] R. B. Randall, *Vibration-based condition monitoring: industrial, aerospace and automotive applications*. John Wiley & Sons, 2011, publication Title: Mechanisms and Machine Science ISSN: 22110992.

[3] S. AG, “Readiness for predictive maintenance at scale report 2023,” Erlangen, Germany, Tech. Rep., 2023.

[4] Y. Lei, B. Yang, X. Jiang, F. Jia, N. Li, and A. K. Nandi, “Applications of machine learning to machine fault diagnosis: A review and roadmap,” *Mechanical Systems and Signal Processing*, vol. 138, p. 106587, 2020, publisher: Elsevier Ltd. [Online]. Available: https://doi.org/10.1016/j.ymssp.2019.106587

[5] O. Surucu, S. A. Gadsden, and J. Yawney, “Condition Monitoring using Machine Learning: A Review of Theory, Applications, and Recent Advances,” *Expert Systems with Applications*, vol. 221, p. 119738, Jul. 2023. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0957417423002397

[6] E. Brusa, C. Delprete, and L. G. Di Maggio, “Eigen-spectrograms: An interpretable feature space for bearing fault diagnosis based on artificial intelligence and image processing,” *Mechanics of Advanced Materials and Structures*, pp. 1–13, Jul. 2022. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/15376494.2022.2102274

[7] M. Umar, M. F. Siddique, N. Ullah, and J.-M. Kim, “Milling Machine Fault Diagnosis Using Acoustic Emission and Hybrid Deep Learning with Feature Optimization,” *Applied Sciences*, vol. 14, no. 22, p. 10404, Nov. 2024. [Online]. Available: https://www.mdpi.com/2076-3417/14/22/10404

[8] A. Mrabti, N. Ouelaa, R. Younes, M. C. Djamaa, and Z. Ouelaa, “Diagnosis of faults using a new VMD-WMRA approach, optimized by a new criterion: Industrial application,” *Advances in Mechanical Engineering*, vol. 16, no. 11, p. 16878132241298371, Nov. 2024. [Online]. Available: https://journals.sagepub.com/doi/10.1177/16878132241298371

[9] F. De Fabritiis and K. Gryllias, “Self-supervised Learning Approach for Anomaly Detection in Rotating Machinery,” *Annual Conference of the PHM Society*, vol. 16, no. 1, Nov. 2024. [Online]. Available: http://www.papers.phmsociety.org/index.php/phmconf/article/view/3911

[10] S. V, T. K, V. C, V. S. D, T. R, and N. A. Patil, “Anomaly Detection in Rotary Equipment using hybrid model with HGSO Optimization,” in *2024 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)*. Bangalore, India: IEEE, Jul. 2024, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/10677203/

[11] R. K. Mishra, A. Choudhary, S. Fatima, A. R. Mohanty, and B. K. Panigrahi, “Realistic Condition-Based Anomaly Detection of Multi-Faults in Rotating Machines,” in *2023 IEEE 3rd International Conference on Sustainable Energy and Future Electric Transportation (SEFET)*. Bhubaneswar, India: IEEE, Aug. 2023, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/10245372/

[12] S. Asavalertpalakorn, P. Singhatanadgid, and T. Ardsomang, “Novelty Detection of a Rolling Bearing using Long Short-Term Memory Autoencoder,” in *2022 37th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC)*. Phuket, Thailand: IEEE, Jul. 2022, pp. 1–4. [Online]. Available: https://ieeexplore.ieee.org/document/9894926/

[13] “CWRU Bearing Data Center.” [Online]. Available: https://engineering.case.edu/bearingdatacenter

[14] J. Hendriks, P. Dumond, and D. Knox, “Towards better benchmarking using the CWRU bearing fault dataset,” *Mechanical Systems and Signal Processing*, vol. 169, p. 108732, Apr. 2022. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0888327021010499

[15] International Organization for Standardization, “Mechanical vibration — Measurement and evaluation of machine vibration — Part 1: General guidelines,” 2016.
[16] ——, “Mechanical vibration — Measurement and evaluation of machine vibration — Part 3: Industrial machinery with a power rating above 15 kW and operating speeds between 120 r/min and 30 000 r/min,” 2022.
[17] C. Delprete and C. Gastaldi, “On the effectiveness of standards application to threshold setting in vibration condition monitoring in industrial machinery,” *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, p. 1748006X231218363, Jan. 2024. [Online]. Available: https://journals.sagepub.com/doi/10.1177/1748006X231218363
[18] E. Brusa, C. Delprete, L. Giorio, L. G. Di Maggio, and V. Zanella, “Design of an Innovative Test Rig for Industrial Bearing Monitoring with Self-Balancing Layout,” *Machines*, vol. 10, no. 54, 2022.
[19] L. G. Di Maggio, L. Giorio, C. Delprete, and E. Brusa, “Dataset of Vibration, Temperature and Speed Measurements for Multiple Types of Localized Defects on Spherical Roller Bearings across Multiple Operating Conditions,” Oct. 2024. [Online]. Available: https://zenodo.org/doi/10.5281/zenodo.13913254
[20] L. G. Di Maggio, E. Brusa, and C. Delprete, “Zero-Shot Generative AI for Rotating Machinery Fault Diagnosis: Synthesizing Highly Realistic Training Data via Cycle-Consistent Adversarial Networks,” *Applied Sciences*, vol. 13, no. 22, p. 12458, Nov. 2023. [Online]. Available: https://www.mdpi.com/2076-3417/13/22/12458
[21] L. G. Di Maggio, “Intelligent Fault Diagnosis of Industrial Bearings Using Transfer Learning and CNNs Pre-Trained for Audio Classification,” *Sensors*, vol. 23, no. 1, p. 211, Dec. 2022. [Online]. Available: https://www.mdpi.com/1424-8220/23/1/211
[22] S. L. Shuo Liu, X. D. Xuemei Ding, and D. C. Damien Coyle, “Integrated Autoencoder-Level Set Method Outperforms Autoencoder for Novelty Detection,” in *2022 International Joint Conference on Neural Networks (IJCNN)*. Padua, Italy: IEEE, Jul. 2022, pp. 1–8. [Online]. Available: https://ieeexplore.ieee.org/document/9891877/
[23] A. Fernandes, V. Leithardt, and J. F. Santana, “Novelty detection algorithms to help identify abnormal activities in the daily lives of elderly people,” *IEEE Latin America Transactions*, vol. 22, no. 3, pp. 195–203, Mar. 2024. [Online]. Available: https://ieeexplore.ieee.org/document/10431423/
[24] F. T. Liu, K. M. Ting, and Z.-H. Zhou, “Isolation Forest,” in *2008 Eighth IEEE International Conference on Data Mining*. Pisa, Italy: IEEE, Dec. 2008, pp. 413–422. [Online]. Available: http://ieeexplore.ieee.org/document/4781136/
[25] C. Cortes and V. Vapnik, “Support-vector networks,” *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
[26] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, “LOF: identifying density-based local outliers,” in *Proceedings of the 2000 ACM SIGMOD international conference on Management of data*. Dallas Texas USA: ACM, May 2000, pp. 93–104. [Online]. Available: https://dl.acm.org/doi/10.1145/342009.335388
[27] M. Ashrafuzzaman, S. Das, A. A. Jillepalli, Y. Chakhchoukh, and F. T. Sheldon, “Elliptic Envelope Based Detection of Stealthy False Data Injection Attacks in Smart Grid Control Systems,” in *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*. Canberra, ACT, Australia: IEEE, Dec. 2020, pp. 1131–1137. [Online]. Available: https://ieeexplore.ieee.org/document/9308523/