Anomaly Detection Techniques in the Service of Data Labeling for Fault Diagnosis in Manufaturing

Aldonso Martins de O. Junior[1,2][a], Emmanuel A. de B. Santos[1][b], Denis Leite[1,3][c] and Alexandre M. A. Maciel[3][d]

[1] *Universidade de Pernambuco, Brazil*

[2] *Stellantis, Brazil*

[3] *Mekatronik I.C. Automacao Ltda, Brazil*

Keywords: Unsupervised Anomaly Detection, Anomaly Detection Models, Fault Identification, Fault Detection, Labeling Process.

Abstract: The lack of labeled fault data in industrial environments presents a major challenge for developing effective fault detection and diagnosis models. This study investigates the application of unsupervised anomaly detection techniques to identify abnormal machine behavior without relying on labeled data. By enabling the early detection of anomalous conditions, these techniques assist in distinguishing normal from faulty instances, supporting the labeling process for improved fault diagnosis. Ten different techniques are evaluated across multiple performance metrics to determine their effectiveness in industrial fault detection. Experimental results demonstrate that Angle-Based Outlier Detection (ABOD) outperformed other methods, achieving a higher F1-score and improved accuracy in recognizing unseen normal data. These findings highlight the potential of unsupervised learning for enhancing industrial fault detection, facilitating the transition to data-driven maintenance strategies, and optimizing data collection processes. The study provides valuable insights into model selection, dataset structuring, and cost-efficient implementation strategies for industrial applications, contributing to the broader adoption of anomaly detection in manufacturing environments.

1 INTRODUCTION

Industry 4.0 is transforming manufacturing and process industries through digital technologies, automation, and data-driven approaches. This shift enhances efficiency, flexibility, and intelligence in production, supply chain, logistics, and maintenance. Key technologies such as machine learning, AI, and IoT drive innovation, necessitating the alignment of machinery and legacy systems with modern standards. To remain competitive, industries integrate these advancements to optimize maintenance and improve operational performance (Ramesh et al., 2020; Dalenogare et al., 2018; Ahmad and Kamaruddin, 2012; Zonta et al., 2020; Tsui et al., 2015; Sakib and Wuest, 2018).

A critical challenge in adopting these technologies for industrial fault detection is the lack of labeled data in the early stages of digital transformation. In many industrial environments, labeled fault data is scarce, poor in quality, or non-existent, complicating the development and training of traditional supervised machine learning models. This limitation hinders the widespread adoption of intelligent fault detection systems, delaying improvements in operational reliability and efficiency.

To address this issue, this study explores the use of unsupervised anomaly detection techniques to differentiate normal and anomalous machine behavior without relying on labeled data. These techniques enable industrial professionals to assess and classify anomalies, refining fault detection and diagnosis processes. By streamlining the labeling of abnormal conditions, this approach enhances the development of more accurate and robust fault detection models.

This study evaluates ten anomaly detection techniques to identify the most effective models for industrial fault detection. Their performance is analyzed across multiple metrics to provide insights into real-world applicability. The hypothesis posits that

[a] https://orcid.org/0000-0001-6479-264X
[b] https://orcid.org/0000-0002-1786-0934
[c] https://orcid.org/0000-0002-0392-3279
[d] https://orcid.org/0000-0003-4348-9291

O. Junior, A. M., Santos, E. A. B., Leite, D. and Maciel, A. M. A.
Anomaly Detection Techniques in the Service of Data Labeling for Fault Diagnosis in Manufaturing.
DOI: 10.5220/0013437100003929
In *Proceedings of the 27th International Conference on Enterprise Information Systems (ICEIS 2025) - Volume 1*, pages 921-928
ISBN: 978-989-758-749-8; ISSN: 2184-4992

unsupervised techniques can detect faults using only normal operational data, addressing the challenge of scarce labeled datasets (Leite et al., 2025). To validate this, a structured experimental approach is implemented, assessing model performance under different dataset configurations to highlight their strengths and limitations in industrial applications.

2 MATERIALS AND METHODS

2.1 Data Sources and Collect

Due to the scarcity of real industrial production datasets (Kang et al., 2020) and the lack of time series data from Digital Manufacturing Machines (DMMs) with both digital and analog IO signals (Leite et al., 2022), simulated data was utilized. This approach, was validated by Huang et al. (Huang et al., 2022) for fault detection, and in this study contemplates two simulators: a pick-and-place system and an electric furnace (Figure 1). They were selected for their contrasting dynamic characteristics, providing distinct profiles to test anomaly detection techniques in industrial settings.

The simulations were run in Unity 3D, a game engine with graphic and physics simulation capabilities. The pick-and-place machine simulated motor forces, friction, and loads, while the electric furnace implemented a dynamic heating model and discrete simulations for door conditions. Each simulated machine was designed to reflect its operational characteristics and potential failure conditions.

- Pick and Place Robot: Sequential machine with linear movement system and short-time cycle. Simulated forces, friction, and loads.
- Industrial Electric Oven: Thermal threatment system, with slow heating and cooling process. Simulated heat transfer, resistance heating system, and door conditions.

2.2 Data Detailing and Simulation

The Pick and Place machine simulation consists on a 3-axis linear positioning system, which operates in a sequential patten of three positions. Two types of faults were simulated for each axis: punctual obstructions (emulating damages like linear guide or fuse issues) and speed losses (representing motor driver power loss, maladjustment, or increased friction). And the corresponding dataset includes 308 normal cycles operation and 103 cycles with each fault type, and the simulation includes the following variables for the Pick and Place system: target position of the three axes, current position of the three axes (analog data), forward and backward commands for the 3 axis.

The Furnace machine simulation features an electric heating system, a temperature sensor, and a door mechanism. It comprises a thermal threating process with heating and colling steps. For this simulation, the following variables are considered: door open signal, door closed signal, heating turned on, maximum power applied, temperature. And the corresponding dataset includes 104 normal cycles and 104 cycles for each fault condition, amounting to a total of 418 cycles. Simulated faults are power loss at the heater, thermal noise, and temperature spam error where included.

2.3 Methodological Structure

This research employs a methodological framework designed to assess the effectiveness of anomaly detection techniques in industrial fault detection. The study is structured into two distinct rounds of experiments, each addressing different aspects of anomaly detection under varying conditions. This approach is illustrated through two block diagrams, which offer a macro perspective of both the experimental setup and the detailed model development and evaluation process.

- 1st Round: utilizes the same data organization as employed by Leite et al. (2022)[8];
- 2nd Round: introduces a new dataset configuration with a varying sampling rate per cycle.

Figure 2 presents the overall framework of the study, beginning with the raw dataset and diverging into two separate pathways for the respective experimental rounds. Round 1 utilizes the dataset organization proposed by Denis et al. (2022) (Leite et al., 2022), focusing on a custom mixed discrete and analogical approach. In contrast, Round 2 delves into a novel dataset configuration with varied sampling number per cycle [10, 25, 50, 100, 200, 350, 400, 500, and 900]. This bifurcation allows for a multifaceted exploration of anomaly detection. Both paths converge at the critical model development and evaluation process.

Ten different anomaly detection techniques were selected to be exploited as lited in Table 1).

The dataset was divided into normal only dataset and faulty dataset (to emulate the reality were the data collection just started in a industry). The normal only dataset is then split in a 70-30 rate, creating the train/test dataset, and the evaluation dataset (here called unseen dataset). The train/test pass through 30

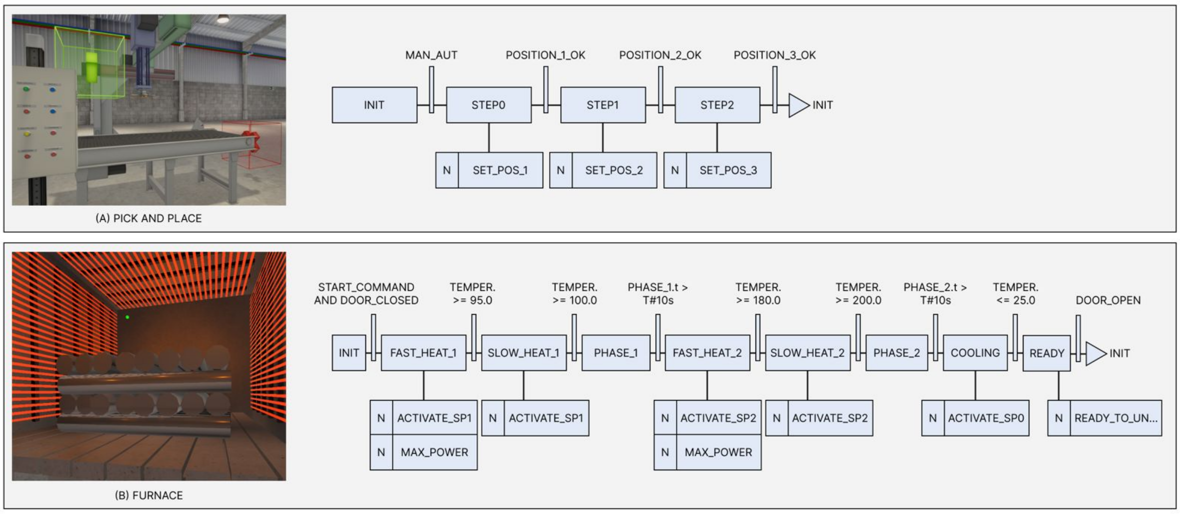

Figure 1: Furnace and Pick and Place Machines Cycle Detailed from Leite et al. (2022) (Leite et al., 2022).

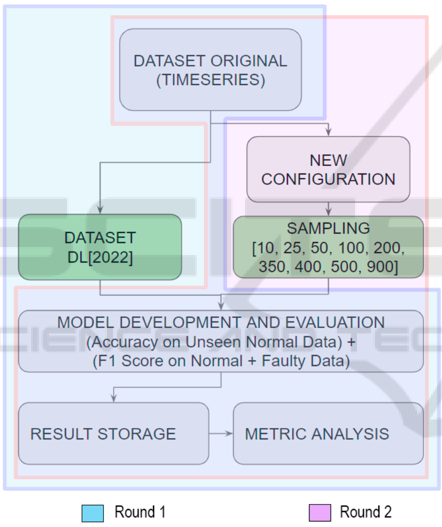

Figure 2: Experimental Setup on Block Diagram with each round conducted.

executions of a new split of 70-30 rate. This time the 70% portion is used to train each unsupervised learning method (Table 1) and the model is used to evaluate it's Accuracy, Recall and F1 Scoring against the test, evaluation (unseen), faulty only and full datasets. All results are recorded for a complete evaluation of the experiments.

The PyCaret low-code library (Ali, 2020) played a central role in the pre-processing and model development process, significantly streamlining key tasks such as feature scaling, train/test split management, outlier removal, and feature selection. Incorporating Sklearn classifiers (Buitinck et al., 2013), PyCaret also facilitated the tuning and optimization of anomaly detection models to achieve the best possible performance. The anomaly detection techniques used in this study, listed in Table 1, were selected based on literature reviews (Albuquerque Filho et al., 2022). To ensure robustness, each technique was used to generate 30 model variations with different random states, mitigating biases from initialization and data splits. Model evaluation followed a two-step process: an initial assessment on test and unseen normal data to measure accuracy in recognizing normal operations, followed by testing on faulty data, where the F1 score served as the primary metric due to the dataset's unbalanced nature.

Table 1: All techniques used on Anomaly Detection Model Creation.

Abbreviation	Description
ABOD	Angle-base Outlier Detection
CLUSTER	Clustering-Based Local Outlier
COF	Connectivity-Based Local Outlier
IFOREST	Isolation Forest
HISTOGRAM	Histogram-based Outlier Detection
KNN	K-Nearest Neighbors Detector
LOF	Local Outlier Factor
SVM	One-class SVM Detector
PCA	Principal Component Analysis
MCD	Minimum Covariance Determinant

As described, the models are trained exclusively with normal data and evaluated on both normal, faulty data and complete Dataset.

2.4 Dataset Preparation and Experimental Details

Round 1 dataset preparation process followed Leite *et al.* (2022). This approach, aims to capture the machine behavior in a specific manner, requiring human contribution only in automation and maintenance domains, and no human contribution in the machine learning domain.

The outcome is a feature set that amalgamates these discrete events with continuous variables, as ilustrated in Figure 3, where each instance in the dataset represents one machine cycle, filled with timed delays of discrete events and the corresponding values of continuous variables. This structure allows for the detection of anomalies such as events occurring out of order, delayed, or early, and issues with calibration, utilities, or machine components:

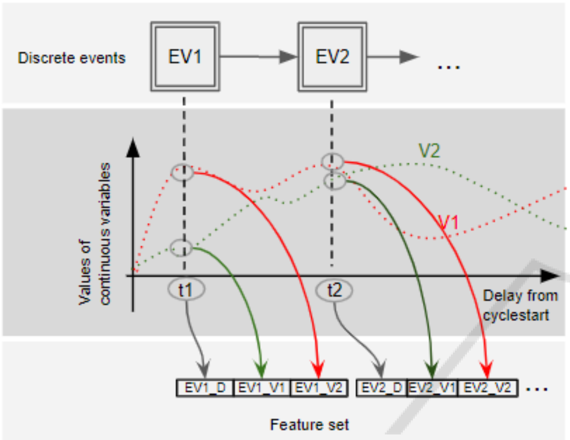

Figure 3: Feature set preparation process combining discrete events and continuous variables.

Round 2 embarked on a more abroad exploration of anomaly detection capabilities by employing a distinct approach to data organization compared. While both rounds focus on datasets that are cycle-oriented, with each instance representing an entire machine cycle, the key difference in Round 2 lies in a new method of feature dataset definition. In this round, the raw time-series data were restructured to capture samples of each variable at a fixed interval within each cycle. This methodology requires less pre-processing effort. However, depending on the number of samples per cycle, it may incur a higher computational cost and higher dimensionality, for the anomaly detection models.

The expectation behind this schema was that a more detailed data granularity would enable a more accurate characterization of the machine's behavior, potentially leading to improved anomaly detection results compared to Round 1. To test this hypothesis, the study examined how different number of samples per cycle, representing a variety of operational scenarios, would affect the performance of anomaly detection models. This involved creating various versions of the dataset with different sampling intervals. Figure 4 shows an example of how the main variables from the Furnace machine would look at different using 2 sampling rates.

Table 2: Different sample sizes for both experiments on Round 2.

samples	Furnace	PickandPlace
10	X	X
25	X	X
50	X	X
100	X	X
200	X	X
350		X
400	X	
500	X	
900	X	

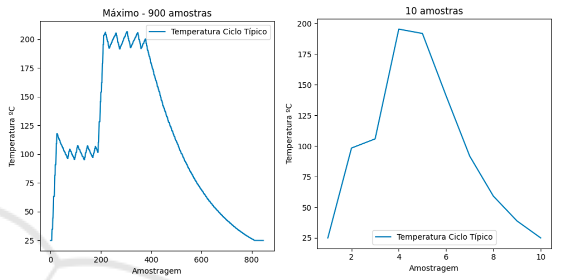

Figure 4: Illustration of feature temperature in Furnace with 2 differents sample sizes.

3 RESULTS

As outlined in the Materials and Methods section, Round 1 involved applying 30 models across 10 selected techniques for each simulated machine, totaling 600 experiments. In Round 2, 8 sampling variants were tested for the Furnace and 6 for the Pick-and-Place machine, as detailed in Table 2, resulting in 2400 models for the Furnace and 1800 for the Pick-and-Place, bringing the total number of evaluated Anomaly Detection (AD) models to 4800. This section presents and discusses the results of both experimental rounds, as illustrated in Figure 2, focusing on accuracy for unseen normal data and the F1 score on the complete dataset (faulty + normal cases).

3.1 Round 1 - Analysis

The Figure 2 highlight the round 1 experiment from which results are going to be discussed on this section. The methodologies and evaluation metrics were consistent with previous descriptions and methodology.

The accuracy assessment on unseen normal (evaluation) data revealed commendable performance across the anomaly detection techniques (see Figure 5 for details). When predicting both faulty and nor-

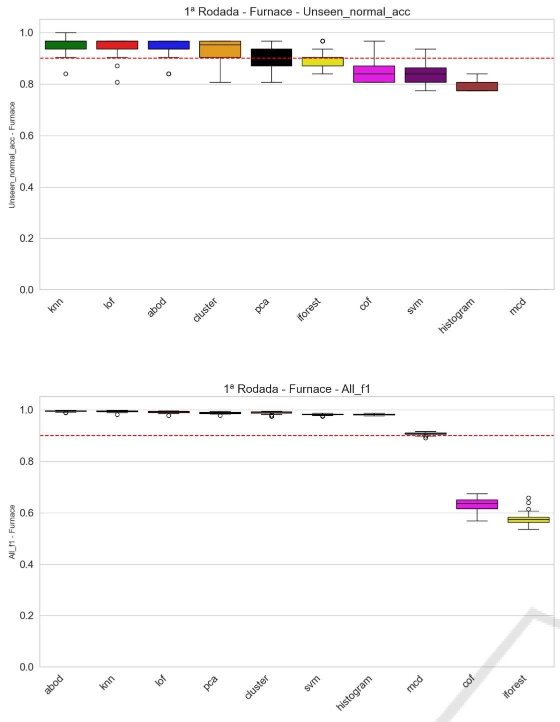

Figure 5: Results Accuracy on Unseen Normal Data + F1 Score on All Data - Furnace.

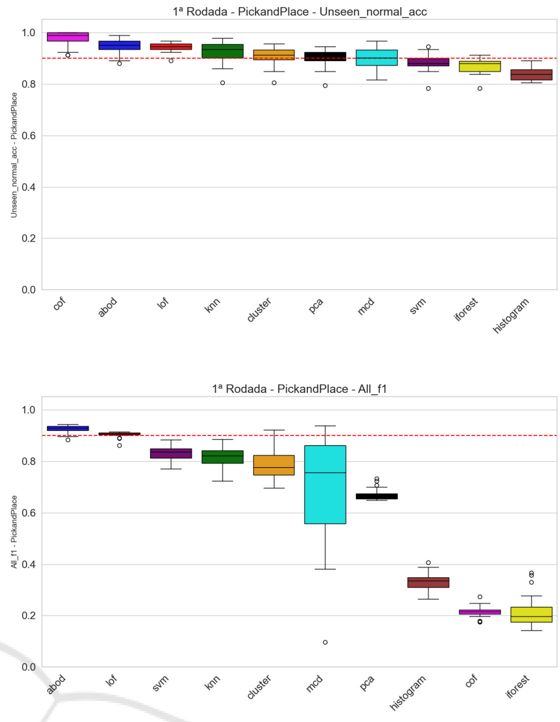

Figure 6: Results Accuracy on Unseen Normal Data + F1 Score on All Data - Pick and Place.

mal (all data set), the F1 score provided insights into the overall effectiveness of the models. All of them demonstrated robust performance (above 0.9 of F1) except for COF and IFOREST which exhibited varying levels of effectiveness (see Figure 5 for details) below 0.7.

The qualitative comparison emphasized the strong performance of ABOD, KNN, LOF, PCA, CLUSTER, SVM, and histogram in terms of accuracy on unseen normal data and F1 score on faulty plus normal data, demonstrating a high ability to distinguish between normal and faulty conditions.

The same methodologies were applied to evaluate anomaly detection techniques by extending the investigation to the pick-and-place machine. The accuracy of unseen normal data revealed distinct patterns of performance for the pick-and-place machine.

When predicting both faulty and normal data, ABOD, LOF, and SVM consistently outperformed other techniques, effectively distinguishing between normal and faulty conditions, whereas Histogram, COF, and IFOREST exhibited lower effectiveness (see Figure 6). Although COF achieved high accuracy on unseen normal data for the Pick-and-Place machine, it failed to generalize for fault detection, showing a significant drop in F1 score when faulty data was introduced. In contrast, ABOD maintained strong performance across both normal accuracy and F1 score, reinforcing its reliability. MCD exhibited unstable results, with noticeable variations in F1 scores, while Histogram, COF, and IFOREST struggled with fault identification. A holistic comparison between the Furnace and Pick-and-Place experiments reveals distinct patterns in anomaly detection performance, with ABOD emerging as the most consistently effective technique across both systems, demonstrating its versatility in different industrial contexts.

3.2 Round 2 Analysis

As a similar approach (but more diverse), the Figure 2 highlights the round 2 experiments and the it's results to be discussed on this section.

The experiments on the Furnace machine provided intriguing insights into the performance of anomaly detection techniques across different sample sizes as show in Figure 7. Notably, as the sample size increased, there was a general trend of improvement in both Unseen Normal Accuracy and All F1 Score. ABOD consistently secured a place in the top four performers for Unseen Normal Accuracy in seven out of eight experiments (see Table 5), showcasing robustness across diverse sample sizes.

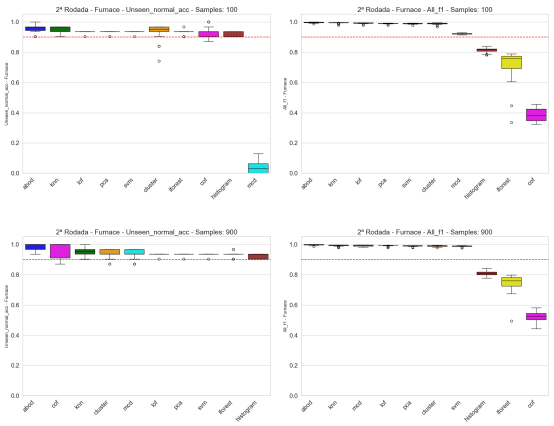

Figure 7: Results for Furnace on different samples [100 and 900 columns].

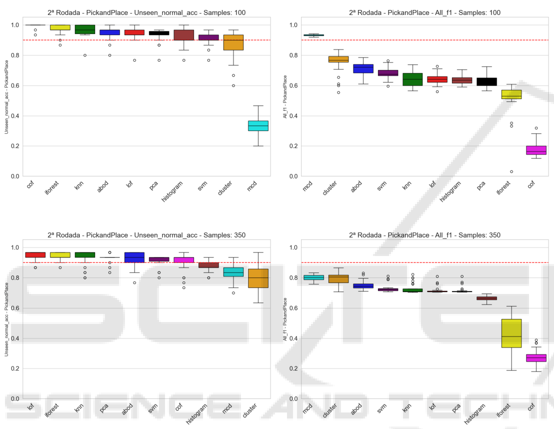

Figure 8: Results for Pick and Place on different samples [100 and 350 columns].

KNN and LOF also demonstrated notable consistency and effectiveness in Furnace experiments, maintaining high rankings across different sample sizes. This reliability suggests that these techniques might exhibit consistent performance in fault detection within the context of the Furnace machine.

Similar to the Furnace experiments, the Pick and Place machine experiments displayed variations in technique performance (Figure 8 with different sample sizes. ABOD, again, stood out by securing a place in the top four for Unseen Normal Accuracy in four out of six experiments (see Table 5), defying the anticipated variability based on sample size. This consistent performance underscores ABOD's potential applicability in diverse industrial scenarios.

3.2.1 Round 2: Cross-Machine Comparison

The analysis of results from Round 2 experiments across both the Furnace and Pick and Place machines, as show in Tables 3 and 4 (in green are the best for each group), has highlighted the remarkable and consistent performance of ABOD. This is particularly striking in the Furnace experiments, where ABOD's effectiveness was notably pronounced. Such findings gain prominence considering ABOD's absence in recent literature reviews on anomaly detection in industrial settings (Albuquerque Filho et al., 2022). This oversight in the literature emphasizes the novelty of our results, suggesting that ABOD may have unique and previously unexplored capabilities for anomaly detection in certain industrial environments.

Table 3: Results for Unseen Normal Accuracy (samples x techniques x machines).

tipo_problema	technique	10	25	50	100	200	350	400	500	900
Furnace	abod	0.924	0.932	0.937	0.962	0.972		0.978	0.944	0.985
Furnace	cluster	0.879	0.913	0.929	0.934	0.943		0.947	0.906	0.943
Furnace	cof	0.908	0.968	0.947	0.927	0.952		0.962	0.910	0.963
Furnace	histogram	0.876	0.923	0.914	0.923	0.924		0.923	0.852	0.923
Furnace	iforest	0.903	0.933	0.935	0.933	0.935		0.932	0.884	0.934
Furnace	knn	0.940	0.935	0.944	0.953	0.953		0.953	0.925	0.954
Furnace	lof	0.930	0.938	0.934	0.934	0.934		0.934	0.909	0.934
Furnace	mcd	0.959	0.000	0.000	0.039	0.616		0.887	0.913	0.937
Furnace	pca	0.895	0.933	0.934	0.934	0.934		0.934	0.904	0.934
Furnace	svm	0.894	0.928	0.933	0.934	0.934		0.934	0.885	0.934
PickandPlace	abod	0.962	0.974	0.947	0.949	0.926	0.926			
PickandPlace	cluster	0.892	0.890	0.848	0.859	0.802	0.802			
PickandPlace	cof	1.000	0.990	0.990	0.997	0.906	0.906			
PickandPlace	histogram	0.860	0.818	0.839	0.913	0.880	0.880			
PickandPlace	iforest	0.972	0.922	0.956	0.968	0.943	0.944			
PickandPlace	knn	0.960	0.937	0.937	0.963	0.934	0.934			
PickandPlace	lof	0.941	0.941	0.920	0.940	0.950	0.950			
PickandPlace	mcd	0.963	0.000	0.017	0.330	0.847	0.847			
PickandPlace	pca	0.926	0.910	0.904	0.930	0.930	0.930			
PickandPlace	svm	0.910	0.893	0.883	0.903	0.917	0.917			

Table 4: Results for All Data F1 Score (samples x techniques x machines).

tipo_problema	technique	10	25	50	100	200	350	400	500	900
Furnace	abod	0.994	0.995	0.995	0.996	0.996		0.997	0.994	0.997
Furnace	cluster	0.985	0.988	0.988	0.988	0.989		0.989	0.987	0.989
Furnace	cof	0.291	0.415	0.457	0.384	0.421		0.446	0.521	0.524
Furnace	histogram	0.819	0.836	0.827	0.814	0.809		0.809	0.794	0.809
Furnace	iforest	0.703	0.734	0.744	0.714	0.726		0.733	0.761	0.745
Furnace	knn	0.993	0.991	0.992	0.993	0.993		0.993	0.990	0.993
Furnace	lof	0.992	0.991	0.992	0.992	0.991		0.991	0.989	0.991
Furnace	mcd	0.989	0.918	0.914	0.921	0.967		0.988	0.991	0.992
Furnace	pca	0.988	0.989	0.989	0.990	0.990		0.990	0.988	0.990
Furnace	svm	0.987	0.988	0.989	0.989	0.989		0.989	0.986	0.989
PickandPlace	abod	0.660	0.650	0.702	0.713	0.750	0.750			
PickandPlace	cluster	0.696	0.705	0.760	0.756	0.794	0.794			
PickandPlace	cof	0.093	0.201	0.131	0.180	0.273	0.273			
PickandPlace	histogram	0.677	0.689	0.683	0.637	0.662	0.662			
PickandPlace	iforest	0.526	0.591	0.576	0.515	0.414	0.414			
PickandPlace	knn	0.613	0.652	0.665	0.644	0.726	0.726			
PickandPlace	lof	0.618	0.648	0.654	0.642	0.717	0.717			
PickandPlace	mcd	0.719	0.967	0.967	0.931	0.800	0.800			
PickandPlace	pca	0.629	0.657	0.651	0.625	0.716	0.716			
PickandPlace	svm	0.660	0.688	0.682	0.682	0.728	0.728			

Increasing the number of samples improved model performance, but without significant changes, suggesting that data collection does not need to occur every 50 ms (the maximum for certain OPC UA hardware). A longer sampling interval can be used without major impacts on results, reducing technological costs and investments in data acquisition software. The consistent performance of ABOD, along with the reliability of KNN and LOF in the Furnace machine, highlights the complexity of anomaly detection effectiveness and the need for further exploration. These findings, rarely addressed in recent literature, provide valuable insights into fault detection methodologies in industrial settings. Notably, IFOREST and

COF performed poorly for both machines when exposed to faulty data, reinforcing their limitations in handling abnormal conditions.

Table 5: Summary of the number of times each technique was in the top 4 for each machine regarding sample sizes experiments.

(technique)	Furnace (unseen_normal_count)	Furnace (all_f1_count)	PickandPlace (unseen_normal_count)	PickandPlace (all_f1_count)
abod	7.0	8.0	4.0	4.0
knn	8.0	8.0	5.0	0.0
lof	3.0	8.0	3.0	0.0
pca	2.0	5.0	2.0	0.0
mcd	2.0	3.0	1.0	6.0
cluster	3.0	0.0	0.0	6.0
cof	6.0	0.0	4.0	0.0
histogram	0.0	0.0	0.0	3.0
iforest	1.0	0.0	5.0	0.0
svm	1.0	0.0	0.0	5.0

4 DISCUSSION

The presented research embarked on a wide exploration of anomaly detection techniques in the context of fault detection for industrial machinery. Through two rounds of investigation, the study examined different techniques across varying sample sizes and two distinct machines: the Furnace and the Pick and Place.

In both experimental rounds, models were initially evaluated using only normal data, reflecting the typical scenario of a new industry implementing Industry 4.0 capabilities. This allowed for selecting the most suitable model before assessing its performance in detecting real faults emulated through full dataset exposure. The first round demonstrated the effectiveness of various anomaly detection techniques in identifying faults in the Furnace and Pick-and-Place machines. ABOD consistently achieved high Unseen Normal Accuracy across different sample sizes, challenging the assumption that technique effectiveness is highly dependent on sample size. Additionally, KNN and LOF proved to be reliable performers in the Furnace machine, reinforcing their potential for robust fault detection.

The second round further explored the impact of sample size on anomaly detection. ABOD continued to deliver stable results across both machines, reinforcing its versatility and questioning conventional assumptions about sample size reliability. A key factor in ABOD's success is its reduced susceptibility to the curse of dimensionality, as noted by Kriegel et al. (Kriegel et al., 2008). By analyzing the variance of angles between data vectors, ABOD effectively identifies outliers, making it a promising candidate for industrial fault detection. This distinctive approach warrants further investigation, potentially paving the way for advancements in anomaly detection methodologies.

This study opens avenues for future investigations, encouraging researchers and practitioners to explore anomaly detection techniques beyond conventional expectations, particularly in scenarios with varying sample sizes. Additionally, it highlights the intricate relationship between sample sizes, machine types, and anomaly detection methods (as shown in Tables 3 and 4). These findings emphasize the need for a nuanced approach where technique selection is guided by established practices and empirical observations tailored to the specific industrial context. Naturally, despite efforts to model and incorporate real-world randomness, studies based on simulated data have inherent limitations. Therefore, future work should explore these techniques using real-world data to further validate and refine the findings.

Furthermore, no comparison with other studies was performed due to the scarcity of benchmark datasets from industrial manufacturing machines, which a key challenges on establishing the state of the art in this fault diagnosis field(Leite et al., 2025), Therefore, new studies may benefit from investigating new techniques over the same dataset used in this study by contacting the authors.

5 CONCLUSIONS

The outcomes of this study significantly address the predicament of limited labeled data in the discrete manufacturing industry. By exploring ten different anomaly detection techniques, trained exclusively with normal data, we have quantified their effectiveness in distinguishing normal from faulty conditions. The results show that the Angle-Based Outlier Detection (ABOD) technique achieved an average increase in F1-score compared to other methods, demonstrating its suitability for real-world applications in fault detection.

One of the most prominent findings is the effectiveness of ABOD in handling manufacturing data, where it consistently outperformed other models. This insight is particularly valuable for researchers and professionals seeking robust and scalable anomaly detection methods for industrial machinery.

The comparative analysis between the two dataset configurations led to a surprising conclusion: the less detailed approach from previous work proved more effective in fault detection than the more comprehensive setup. This finding highlights that the quality and relevance of data are often more critical than sheer quantity when constructing training sets for anomaly detection models.

Ultimately, this research contributes insights to the field of fault detection in industrial systems, challenging conventional notions and paving the way for more nuanced and effective anomaly detection methodologies. The findings encourage a shift toward a more adaptive and context-aware approach in implementing anomaly detection techniques for diverse industrial applications. These insights provide a strong foundation for future studies aimed at refining anomaly detection strategies and validating them in real-world manufacturing settings.

ACKNOWLEDGEMENTS

The authors would like to thank the research team GPCDA (Grupo de Pesquisa em Ciência de Dados e Analytics) at PPGEC-UPE for all the invaluable contributions to the discussions and for their support. We need also to thank the huge contribution of Mekatronik's Team for sharing the simulation data and for the opportunity to work on methodology and experiment execution. And to thank Stellantis Goiana for all the support and incentive. Gratitude is extended to CAPES and CNPq for their financial support, without which this work would not have been possible.

This paper was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES) - Finance Code 001, Fundação de Amparo a Ciência e Tecnologia do Estado de Pernambuco (FACEPE), the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) - Brazilian research agencies.

REFERENCES

Ahmad, R. and Kamaruddin, S. (2012). An overview of time-based and condition-based maintenance in industrial application. *Computers &industrial engineering*.

Albuquerque Filho, J. E. D., Brandão, L. C. P., Fernandes, B. J. T., and Maciel, A. M. A. (2022). A review of neural networks for anomaly detection. *IEEE Access*, 10:112342–112367.

Ali, M. (2020). *PyCaret: An open source, low-code machine learning library in Python*. PyCaret version 2.3.

Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., Niculae, V., Prettenhofer, P., Gramfort, A., Grobler, J., Layton, R., VanderPlas, J., Joly, A., Holt, B., and Varoquaux, G. (2013). API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pages 108–122.

Dalenogare, L., Benitez, G., Ayala, N., and ... (2018). The expected contribution of industry 4.0 technologies for industrial performance. *International Journal of ...*.

Huang, J., Wen, J., Yoon, H., Pradhan, O., Wu, T., and Neill, Z. O. (2022). Energy & Buildings Real vs . simulated : Questions on the capability of simulated datasets on building fault detection for energy efficiency from a data-driven perspective. *Energy & Buildings*, 259:111872.

Kang, Z., Catal, C., and Tekinerdogan, B. (2020). Machine learning applications in production lines : A systematic literature review. *Computers & Industrial Engineering*, 149(August):106773.

Kriegel, H., Schubert, M., and Zimek, A. (2008). Angle-based outlier detection in high-dimensional data. *Proceedings of the 14th ACM ...*.

Leite, D., Andrade, E., Rativa, D., and Maciel, A. M. (2025). Fault Detection and Diagnosis in Industry 4.0: A Review on Challenges and Opportunities. *Sensors*, 25(1):1–34.

Leite, D., Martins, A. J., Rativa, D., Oliveira, J. D., and Maciel, A. (2022). An automated machine learning approach for real-time fault detection and diagnosis. *Sensors*.

Ramesh, P., Dutta, S., Neog, S., Baishya, P., and ... (2020). Implementation of predictive maintenance systems in remotely located process plants under industry 4.0 scenario. ... *Engineering: In Honor of ...*.

Sakib, N. and Wuest, T. (2018). Challenges and opportunities of condition-based predictive maintenance: a review. *Procedia cirp*.

Tsui, K., Chen, N., Zhou, Q., Hai, Y., and Wang, W. (2015). Prognostics and health management: A review on data driven approaches. *Mathematical Problems in ...*.

Zonta, T., Costa, C. D., da Rosa Righi, R., and ... (2020). Predictive maintenance in the industry 4.0: A systematic literature review. *Computers &Industrial ...*.