Predictive Maintenance of Machines in LABtop Production Environment

Primož Kocuvan
Department of intelligent systems
"Jožef Stefan" Institute
Ljubljana, Slovenia
primoz.kocuvan@ijs.si

Vinko Longar
Rudolfovo - znanstveno in tehnološko središče
Novo mesto, Slovenia
vinko.longar@rudolfovo.eu

Rok Struna
Rudolfovo - znanstveno in tehnološko središče
Novo mesto, Slovenia
rok.struna@rudolfovo.eu

Abstract

This study investigates predictive maintenance of CNC machinery within the LABtop production environment through the deployment of iCOMOX sensor modules on a compressor and machine spindle. Each module integrates multi-modal sensing capabilities, including vibration, magnetic field, temperature, and acoustic measurements, enabling comprehensive monitoring of machine conditions. Data was collected at five-minute intervals over a 30-day period, resulting in an unlabeled dataset due to the absence of recorded failures or anomalies. The analysis employed unsupervised machine learning techniques, specifically principal component analysis (PCA) for dimensionality reduction and clustering to identify operational patterns. PCA successfully reduced the original 11-dimensional dataset to two principal components, allowing for effective visualization and grouping. The elbow and silhouette methods determined three optimal clusters for both sensors, with one cluster in each case identified as a potential outlier. Results suggest that dense clusters represent normal operation, while outlier clusters may indicate measurement errors or emerging machine faults. Although supervised learning could not yet be applied, future work will integrate fault-labeled data to enable robust predictive maintenance models.

Keywords

predictive maintenance, PCA method, production environment, silhouette analysis, elbow method.

1 Introduction

The increasing complexity of modern production systems demands advanced approaches to machine maintenance in order to minimize downtime, reduce costs, and ensure consistent product quality. Traditional maintenance strategies, such as corrective or preventive maintenance, often fail to provide early warnings of failures and may result in either excessive servicing or unexpected breakdowns. Predictive maintenance, by contrast, leverages sensor data and machine learning techniques to detect patterns, identify anomalies, and forecast potential failures before they occur. This approach not only enhances operational efficiency but also extends the lifetime of critical equipment.,

Within the LABtop production environment (consists of multiple machines in sequence - mostly drilling and cutting machines), predictive maintenance has been explored through the integration of advanced multi-sensor monitoring solutions. For this purpose, the public research institute *Rudolfovo* implemented iCOMOX sensor modules on both the compressor and the spindle of a CNC machine. Each iCOMOX module integrates several sensing elements—vibration, magnetic field, temperature, and acoustic measurements—providing a rich dataset suitable for machine learning–based condition monitoring.

The collected data were acquired over a continuous 30-day period at five-minute intervals. Since no machine failures, temperature anomalies, or bearing defects were recorded during this time, the dataset lacked diagnostic labels and was therefore treated as unlabeled. To address this, unsupervised learning methods were employed to uncover latent structures in the data. Principal component analysis (PCA) was used to reduce the dimensionality of the dataset, while clustering methods were applied to identify patterns and potential anomalies in machine operation. The aim of this study is to evaluate the feasibility of unsupervised learning methods in predictive maintenance for industrial equipment, specifically under conditions where fault-labeled data are unavailable. By analyzing the clustering behavior of sensor signals, this work provides insights into normal operating regimes and potential deviations that may correspond to early indicators of faults or measurement errors. Future work will incorporate supervised learning techniques once labeled fault data become available, enabling the development of robust predictive models.

http://doi.org/10.70314/is.2025.skui.0545

2 Related Work

The field of predictive maintenance (PdM) has advanced considerably, with strong emphasis on unsupervised learning methods for anomaly detection and health assessment when labeled failure data are unavailable. PdM has been shown to significantly reduce maintenance costs, decrease unexpected downtime, and enhance equipment reliability [1]. Multi-sensor monitoring platforms such as iCOMOX have emerged as versatile tools for industrial condition monitoring. These devices integrate vibration, magnetic field, temperature, and acoustic sensors into a compact, industrial-grade package capable of edge analytics and cloud integration [2–5].

Such systems enable continuous monitoring of machine health and support the implementation of predictive maintenance strategies in Industry 4.0 environments. From a methodological perspective, unsupervised learning techniques, such as principal component analysis (PCA) and clustering, are widely applied for exploratory data analysis, dimensionality reduction, and anomaly detection. A comprehensive survey highlights the breadth and maturity of these techniques across domains [6]. Clustering methods including k-means, DBSCAN, and OPTICS are instrumental in grouping operational states and unveiling deviations that may signify incipient failures [7].

Hybrid methods combining PCA with clustering have proven effective in enhancing fault detection capabilities. For example, a railcar health monitoring system employing DBSCAN clustering with PCA achieved fault detection accuracy of 96.4% [8]. Similarly, kernel PCA has been applied to construct health indices for unsupervised prognostics [9]. In compressor maintenance, incorporating clustering-derived features into supervised classifiers improved predictive accuracy by 4.9% and reduced training time by 23% [10]. Several studies also propose frameworks that integrate unsupervised learning with IoT and Big Data infrastructures, enabling scalable predictive maintenance solutions across industrial environments [11]. These works demonstrate the feasibility of extracting actionable health indicators from unlabeled sensor data and underscore the critical role of advanced analytics in industrial condition monitoring.

3 Methodology

3.1 Data Acquisition

Two iCOMOX sensor modules were installed on critical machine components within the LABtop production system: the spindle of a CNC machine and the air compressor. Each sensor module integrates vibration, magnetic field, temperature, and acoustic sensing elements, thereby providing multimodal monitoring capabilities. Data were sampled at 5-minute intervals over a continuous 30-day observation period, resulting in an unlabeled dataset due to the absence of recorded failures, anomalies, or maintenance events.

3.2 Data Preprocessing

Raw signals from the iCOMOX modules were aggregated into feature vectors, yielding an 11-dimensional dataset. Standard preprocessing steps included: normalization of features to remove scaling effects, filtering to reduce noise (particularly in the acoustic and vibration signals), and synchronization of multimodal sensor streams.

3.3 Dimensionality Reduction

To facilitate visualization and clustering, dimensionality reduction was performed. Multiple techniques (e.g., t-SNE, Isomap, and autoencoders) were evaluated; however, Principal Component Analysis (PCA) demonstrated superior stability and interpretability. The data were reduced from 11 to 2 principal components, which captured the majority of the variance and allowed effective 2D representation.

3.4 Clustering Analysis

Clustering was applied to the reduced dataset to uncover hidden structures and potential anomalies. The elbow method (*Figure 1*) and silhouette coefficient (*Figure 2*) were employed to determine the optimal number of clusters. Based on these metrics, three clusters were identified for each sensor dataset.

The analysis was conducted separately for the two sensor modules (iCOMOX1 on the spindle and iCOMOX2 on the compressor). Outlier clusters were identified and highlighted for subsequent interpretation.

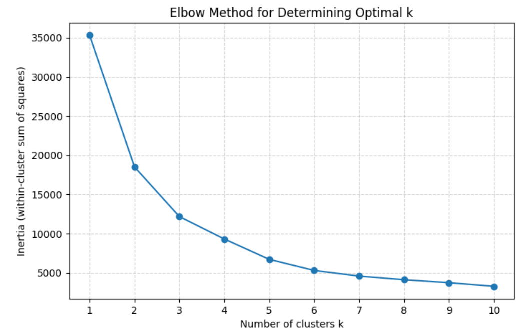

Figure 1: Elbow method

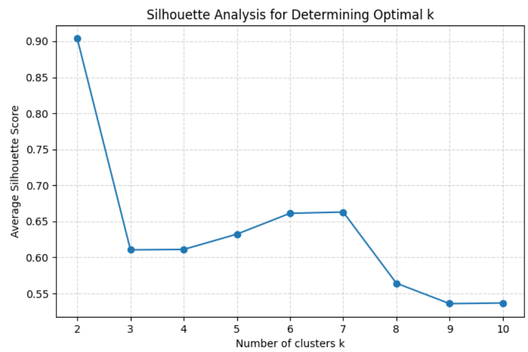

Figure 2: Silhouette coefficient

4 Results

PCA successfully compressed the 11-dimensional dataset into a two-dimensional space. The first two principal components explained the majority of the variance (>80%), enabling effective visualization of patterns in machine behavior. *Figure 3* illustrates the scatter plot for iCOMOX1. Three distinct clusters are visible, with Cluster 1 (highlighted in orange) showing divergence from the main operating regime. *Figure 4* presents the scatter plot for iCOMOX2, where Cluster 2 (highlighted in green) emerges as an outlier relative to the normal operating clusters.

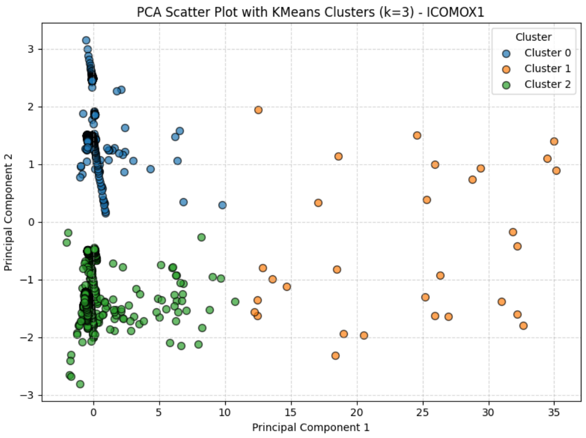

Figure 3: PCA Scatter plot for ICOMOX1

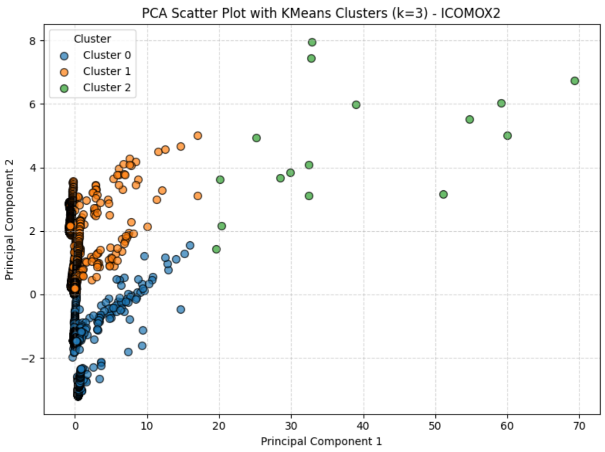

Figure 4: PCA Scatter plot for ICOMOX2

Clusters containing densely grouped points correspond to normal operating conditions of the CNC spindle and compressor. The outlier clusters, however, represent either:

- sensor noise or measurement anomalies (e.g., transient vibration spikes or acoustic distortions), or
- incipient machine faults, which could not be conclusively confirmed due to the absence of ground-truth failure data.

- PCA combined with clustering effectively distinguished between normal operation and anomalous behavior.
- Both sensor datasets (iCOMOX1 and iCOMOX2) revealed three clusters, with one consistently standing out as an outlier.
- Without diagnostic labels, these outliers cannot be definitively classified as machine faults, but their presence highlights potential events of interest for further investigation.
- The results validate the feasibility of unsupervised learning for predictive maintenance in environments lacking labeled fault data.

5 Discussion

The findings from this study demonstrate the viability of unsupervised learning methods in particular PCA and clustering for analyzing unlabeled condition-monitoring data in industrial environments. By reducing an 11-dimensional dataset to two principal components, it was possible to visualize operational states and uncover outlier clusters that may correspond to anomalous machine behavior. This outcome aligns with previous work emphasizing the effectiveness of dimensionality reduction and clustering in predictive maintenance tasks where labeled fault data are limited or unavailable [6,8,9].

The observation of three clusters for both the spindle (iCOMOX1) and compressor (iCOMOX2) highlights the presence of distinct operating regimes within the LABtop system. The fact that one cluster consistently emerged as an outlier suggests potential precursors to faults or, alternatively, sensor-related anomalies. While conclusive interpretation requires diagnostic labels, the clustering nevertheless provides an essential first step toward identifying patterns that can later inform supervised learning models once fault data become available.

Compared to related studies, the present results confirm trends reported in railcar health monitoring [8] and compressor maintenance [10], where unsupervised approaches successfully revealed structural patterns in the absence of labeled datasets. The advantage of PCA lies in its ability to preserve variance while simplifying visualization, which proved more effective than alternative reduction methods considered here (e.g., t-SNE or Isomap). This echoes findings from other industrial applications where PCA has served as a reliable baseline for anomaly detection [9].

An important implication is that multi-sensor platforms such as iCOMOX provide the richness of data required for advanced analytics. The combination of vibration, acoustic, magnetic field, and temperature measurements enables detection of subtle variations that might not be visible through single-sensor monitoring. As highlighted in prior work [2–5], the integration of multimodal data streams significantly strengthens predictive maintenance frameworks by improving robustness and interpretability.

Nevertheless, this study also underscores the limitations of unsupervised learning. Without failure labels, it is not possible to conclusively distinguish between anomalies arising from true machine faults and those caused by sensor noise or environmental conditions. This limitation has been widely noted in the literature [6,11]. Future work should therefore focus on generating labeled datasets through controlled fault injection or long-term monitoring until natural failures occur. Such datasets would enable supervised and hybrid learning approaches, which have shown promise in achieving higher predictive accuracy and more actionable decision support [1,10].

In summary, the present analysis validates the potential of unsupervised learning for predictive maintenance in data-scarce environments. While preliminary, the results establish a methodological foundation for extending condition monitoring at LABtop to more advanced machine learning pipelines, ultimately contributing to early fault detection, reduced downtime, and optimized maintenance planning.

6 Future Work

The present study establishes a foundation for predictive maintenance at LABtop using unsupervised learning methods; however, several directions remain open for further investigation.

First, the absence of diagnostic labels limited this study to exploratory clustering and anomaly detection. Future work will prioritize the collection of labeled datasets through either (i) controlled fault injection experiments on non-critical test equipment or (ii) extended operational monitoring until natural failures occur. The availability of labeled fault data will enable the application of supervised learning and hybrid approaches, combining clustering-derived features with classification models to improve fault detection accuracy and reliability, as demonstrated in recent compressor studies [10]. Second, while PCA provided an effective means of dimensionality reduction, more advanced techniques such as kernel PCA, autoencoders, and variational autoencoders should be investigated. These methods may capture nonlinear relationships in the sensor data that PCA cannot, potentially yielding richer health indicators and more precise separation of operational regimes [9]. Third, the present work focused primarily on offline analysis. Future research should extend to real-time streaming analytics, leveraging the edge-processing capabilities of the iCOMOX platform [2–5]. Deploying online anomaly detection models would allow immediate identification of abnormal conditions and facilitate proactive maintenance decisions.

Fourth, integration with IoT and cloud-based platforms remains a key step toward scalable deployment. By embedding unsupervised learning models into Industry 4.0 architectures, LABtop can benefit from centralized monitoring, cross-machine comparisons, and fleet-level anomaly detection, as highlighted in existing frameworks [11].

Finally, interpretability remains an essential concern. Future efforts will explore explainable AI (XAI) techniques to provide actionable insights into why certain clusters or anomalies are flagged, thereby enhancing operator trust and enabling domain experts to validate and refine the models.

Acknowledgments

The research was supported by DIGITOP project which is funded by Ministry of Higher Education, Science and Innovation of Slovenia, Slovenian Research and Innovation Agency, and EU NextGenerationEU under Grant TN-06-0106. We thank prof. dr. Matjaž Gams for proof-reading the article and mentorship support within DIGITOP project.

References

[1] Abdeldjalil Benhanifia, Zied Ben Cheikh, Paulo Moura Oliveira, Antonio Valente, José Lima. *Systematic review of predictive maintenance practices in the manufacturing sector*. Intelligent Systems with Applications, Volume 26, 2025, Article 200501. ISSN 2667-3053. https://doi.org/10.1016/j.iswa.2025.200501

[2] RS Components, *iCOMOX Intelligent Condition Monitoring Box – Product Datasheet*, 2019. Available: https://docs.rs-online.com/c878/A700000007538369.pdf, Accessed 25.8.2025

[3] EE Times Europe, *Arrow introduces new Shiratech iCOMOX condition-based monitoring products*, 2019. Available: https://www.eetimes.eu/press-releases/arrow-introduces-new-shiratech-icomox-condition-based-monitoring-products/, Accessed 25.8.2025

[4] EBOM, *New Shiratech iCOMOX sensor-to-cloud platform cuts time-to-market for intelligent condition monitoring*, 2019. Available: https://www.ebom.com/new-shiratech-icomox-sensor-to-cloud-platform-cuts-time-to-market-for-intelligent-condition-monitoring/, Accessed 25.8.2025

[5] Sensor+Test, *iCOMOX – Condition Monitoring Box*, 2023. Available:https://www.sensor-test.de/assets/Fairs/2023/ProductNews/PDFs/iCOMOX.pdf, Accessed 25.8.2025

[6] K. Taha, "Semi-supervised and un-supervised clustering: A review and experimental evaluation," *Information Systems*, vol. 114, p. 102178, 2023. doi: 10.1016/j.is.2023.102178

[7] GopenAI Blog, *Predictive maintenance with unsupervised machine learning algorithms*, 2020. Available: (Blog.gopenai.com), Accessed 25.8.2025

[8] M. Ejlali, E. Arian, S. Taghiyeh, K. Chambers, A. H. Sadeghi, D. Cakdi, and R. B. Handfield, "Developing Hybrid Machine Learning Models to Assign Health Score to Railcar Fleets for Optimal Decision Making," *arXiv preprint* arXiv:2301.08877, 2023.

[9] Z. Chen et al., "Health Index Construction Based on Kernel PCA for Equipment Prognostics," *Control Engineering Practice*, vol. 126, 2022.

[10] A. Salazar et al., "Unsupervised Feature Extraction for Compressor Predictive Maintenance Using Clustering and Supervised Learning," *arXiv*, 2024.

[11] Nota, Giancarlo, Nota, Francesco, Toro Lazo, Alonso Nastasia, Michele. (2024). A framework for unsupervised learning and predictive maintenance in Industry 4.0. International Journal of Industrial Engineering and Management. 15. 304-319. 10.24867/IJIEM-2024-4-365.