Report for the Course of Digital Forensic and Biometric

Project: Anomaly Detection

Torres Ramos, Juan Luis and Pardo Rueda, Rafael

January 19, 2026

Abstract

Predictive Maintenance of Industrial Machinery using Deep Autoencoders and Isolation Forest

In factories, it's crucial to catch broken machines before they stop working completely. This project builds an intelligent system that listens to industrial fans and detects faults (like unbalance or clogging) just analyzing their sound. A big challenge in this field is that we have plenty of recordings of healthy machines, but very few recordings of broken ones. To solve this, we used a semi-supervised approach: we taught the AI what normal sounds like, so it could flag anything different as an anomaly.

The process involved two main steps. First, we converted the audio recordings into spectrograms and used them in a Convolutional Autoencoder. This deep neural network acted like a compressor: it was trained to memorize the patterns of a healthy motor and ignore noise. Second, we took the compressed summaries of these sounds and used them into an Isolation Forest algorithm. This second model acted as a judge, looking at the summaries to decide which ones were too strange to be normal.

Our results showed that the Autoencoder successfully learned the signature of a healthy fan. When we tested it with broken machines, the system struggled to recognize them, resulting in high error scores that correctly identified them as anomalies. This proves that we can detect mechanical failures using only healthy data for training, making this a practical solution for real-world maintenance.

Keywords
Anomaly Detection, Predictive Maintenance, Deep Learning, Convolutional Autoencoder (CAE), Isolation Forest, Semi-supervised Learning, MIMII Dataset, Mel-Spectrogram, Industrial Internet of Things (IIoT).

1 Objectives

In factories, machines like industrial fans are critical. When they break unexpectedly, it causes costly downtime. The goal of Predictive Maintenance is to listen to machines and fix them before they fail.

The main challenge is that we have thousands of hours of recordings for healthy machines, but almost no recordings of broken ones. Standard AI needs both to learn.

The main objective is to build a system that can detect broken fans without ever seeing a broken one during training. We achieved this by teaching an AI exactly what a normal fan sounds like, so it naturally flags any strange sound (anomaly) as a problem.

2 The Dataset

We used the **MIMII Dataset** (Sound Dataset for Malfunctioning Industrial Machine Investigation and Inspection) [1].

- **Machine Type:** Industrial Fan.
- **Data Source:** Audio recordings (`.wav`) containing machine sounds and background factory noise.
- **Split:**
  - **Training Data:** 100% Healthy sounds.
  - **Test Data:** A mix of Healthy sounds and Broken sounds (Clogging, Voltage change, Unbalance).

3 Methodology

Our approach converted sound into images, compressed them to extract meaning, and then flagged the outliers using statistical analysis.

3.1 Phase A: Preprocessing (Audio → Images)

Since Deep Learning models often perform better with visual data than raw waveforms, we converted the audio recordings into images. This process involved the following steps:

- **Spectrogram Generation:** We transformed the 3-second audio clips into **Log-Mel Spectrograms**. This creates a "heatmap" where the X-axis represents time, the Y-axis represents frequency (pitch), and the color intensity represents loudness.

- **Normalization:** To ensure stable training, we scaled all pixel values to a range between 0 and 1.

- **Resizing:** Each spectrogram was standardized to a fixed resolution of 128 × 128 pixels.

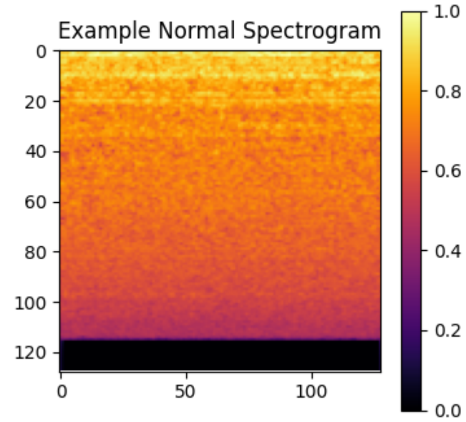

Figure 1: Sample Mel-Spectrogram (Input Data). This visualizes the frequency patterns of a healthy motor.

3.2 Phase B: The Convolutional Autoencoder

To learn the features of a normal motor, we designed and trained a **Convolutional Autoencoder**. This is a specific type of neural network with an "hourglass" architecture, composed of two main parts:

- **The Encoder (Compression):** This section takes the input image (128 × 128) and passes it through a series of convolutional layers. We increased the depth of the network (from 32 to 64 to 128 filters) while progressively reducing the image size using strided convolutions. This forces the model to discard noise and compress the data into a dense, abstract summary known as the **Latent Vector**.

- **The Decoder (Reconstruction):** This section mirrors the encoder. It takes the compressed Latent Vector and uses transpose convolutions (upsampling) to rebuild the original image pixel by pixel.

Training Strategy: The model was trained exclusively on **healthy** data. We used the Mean Squared Error (MSE) loss function to measure the difference between the input and the reconstruction. By minimizing this error, the network was forced to learn the essential acoustic "signature" of a functioning fan (e.g., the rhythmic hum) so it could reconstruct it accurately.

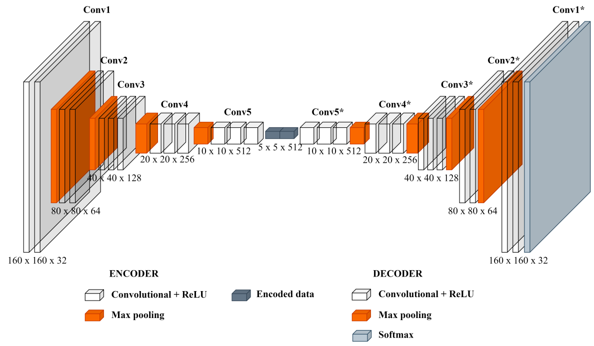

Figure 2: Convolutional Autoencoder Architecture. The input is compressed into a latent vector and then reconstructed [2].

3.3 Phase C: Isolation Forest

In this phase, we apply an *Isolation Forest* (IF) model to perform unsupervised anomaly detection on the latent representations generated by the encoder part of the autoencoder. The goal of this stage is to identify abnormal machine behavior by learning the statistical structure of **healthy operating conditions only**, following a one-class anomaly detection paradigm.

3.3.1 Input Data Representation

Each audio sample is represented by a latent tensor of shape 16 × 16 × 128, produced by the encoder network trained exclusively on healthy sounds. In order to make these representations compatible with classical machine learning algorithms, each tensor is flattened into a one-dimensional feature vector of 32,768 dimensions. This transformation preserves the information content of the latent space while adapting the data format to the requirements of the Isolation Forest algorithm.

Only latent vectors corresponding to healthy samples are used during training. A subset of healthy data is held out for testing, while all available anomalous samples are reserved exclusively for evaluation. This strict separation ensures that the model does not learn any information from faulty conditions during training.

3.3.2 Model Training

The Isolation Forest is trained using only healthy latent vectors. The algorithm operates by recursively partitioning the feature space using randomly selected features and split values. Samples that require fewer splits to be isolated are considered more likely to be anomalous.

Since the training set contains only healthy data, the contamination parameter is set to a very low value, reflecting the assumption that anomalies are rare and should not influence the learned decision boundaries. This configuration allows the model to focus on capturing normal system behavior rather than adapting to outliers.

Once trained, the model outputs:

- a binary prediction for each test sample (normal or anomalous),
- and a decision score that quantifies how strongly a sample deviates from the learned normal pattern.

3.3.3 Evaluation Strategy

The trained Isolation Forest is evaluated on a test set composed of healthy samples unseen during training and all available anomalous samples. Performance is assessed using multiple complementary metrics, including the confusion matrix, precision, recall, and F1-score.

In addition, anomaly score distributions are analyzed to visually inspect the separation between normal and anomalous samples. This multi-level evaluation provides insight into the detector's sensitivity to abnormal conditions as well as its robustness against false alarms.

Overall, this phase establishes a strong statistical baseline for anomaly detection in the learned latent space, which is later compared with alternative approaches such as One-Class SVM to highlight differences in modeling assumptions and detection behavior.

3.4 Phase D: One-Class SVM

In this phase, we evaluate a *One-Class Support Vector Machine* (One-Class SVM) as an alternative statistical method for anomaly detection on the latent representations produced by the autoencoder encoder. The objective of this stage is to analyze how a boundary-based model behaves when learning normal machine operation in the learned latent space and to compare its performance with the Isolation Forest approach described in the previous phase.

3.4.1 Input Data Representation

The input to the One-Class SVM is the same latent representation used in Phase C. Each audio sample is encoded as a latent tensor of shape $16 \times 16 \times 128$, which is subsequently flattened into a one-dimensional feature vector of 32,768 dimensions. This ensures that both anomaly detection models operate on identical feature spaces, enabling a fair comparison.

As in the previous phase, the model is trained exclusively using latent vectors corresponding to healthy operating conditions. Healthy samples not used for training, together with all available anomalous samples, are reserved for evaluation.

3.4.2 Model Training

One-Class SVM learns a decision boundary that encloses the majority of the training samples in feature space. Samples that fall outside this boundary are classified as anomalies. In contrast to Isolation Forest, which isolates samples through random partitioning, One-Class SVM explicitly models the support of the normal data distribution.

Since the model is sensitive to feature scaling, standardization is applied prior to training to ensure that all features contribute equally to the learned boundary. The model is configured with a radial basis function (RBF) kernel, which allows the decision boundary to adapt to non-linear structures in the latent space. The parameter ν controls the fraction of training samples that are allowed to lie outside the learned boundary, reflecting the assumption that anomalies are rare.

3.4.3 Evaluation Strategy

The trained One-Class SVM is evaluated on the same test set used for Isolation Forest, consisting of unseen healthy samples and anomalous samples. Model performance is assessed using confusion matrices and standard classification metrics, including precision, recall, and F1-score.

In addition, anomaly score distributions are analyzed to examine how strongly normal and anomalous samples are separated. This evaluation strategy enables a direct comparison with Isolation Forest and highlights the strengths and limitations of boundary-based anomaly detection in high-dimensional latent spaces.

Overall, this phase provides complementary insight into the suitability of support-based anomaly detection methods for latent representations learned by deep autoencoders, and sets the foundation for a comparative discussion between different statistical approaches.

4 Results

To evaluate the system, we analyzed three key aspects: the stability of the training process, the model's ability to reconstruct healthy sounds, and its accuracy in detecting anomalies.

4.1 Training Performance (Loss Curve)

The first step was to verify that the Autoencoder actually learned. We monitored the Mean Squared Error (MSE), which measures the difference between the original sound and the model's reconstruction.

As shown in Figure 3, the error (Loss) dropped consistently over 30 epochs.

- The **Blue Line (Training)** shows the model learning from the data.
- The **Orange Line (Validation)** follows closely, which is a perfect result. It indicates that the model is not "memorizing" the data (overfitting) but is genuinely learning the general structure of a healthy fan sound.

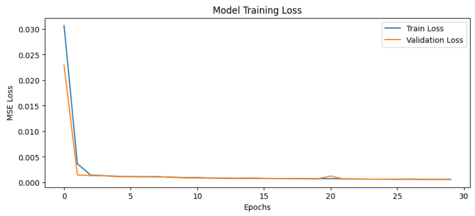

Figure 3: Training and Validation Loss Curve

4.2 Visual Reconstruction Quality

To verify what the model actually learned, we visually compared the original spectrograms (Input) against the Autoencoder's reconstructions (Output). Figure 4 demonstrates this comparison:

- **Top Row (Original):** Shows the clear, detailed frequency patterns of the raw audio.
- **Bottom Row (Reconstructed):** Shows the model's attempt to redraw the sound. While slightly blurrier due to compression, the reconstruction successfully preserves the key structural elements, such as the fundamental frequency and the periodic rhythm of the motor. This confirms that the Latent Space has captured the essential identity of the healthy machine.

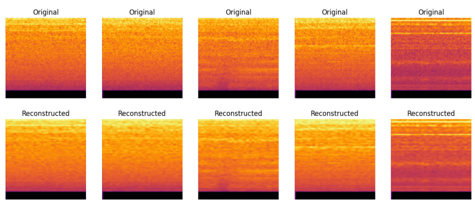

Figure 4: Input vs. Reconstructed Spectrograms

4.3 Anomaly Detection (Error Distribution)

The core of our anomaly detection system relies on the assumption that the model will fail to reconstruct sounds it has never seen (i.e., defects). We calculated the reconstruction error (MSE) for the validation dataset and plotted the distribution in Figure 5.

- **Blue Bars:** Represent the healthy test samples. They cluster tightly on the left side (Low Error), confirming the model recognizes them well.
- **Red Dashed Line:** Represents the anomaly threshold (defined as Mean + 2 Standard Deviations).
- **Analysis:** The histogram shows a clear separation. Healthy data falls below the threshold. In a real-world scenario, any incoming sound with an error falling to the right of this line would be immediately flagged as a potential mechanical failure.

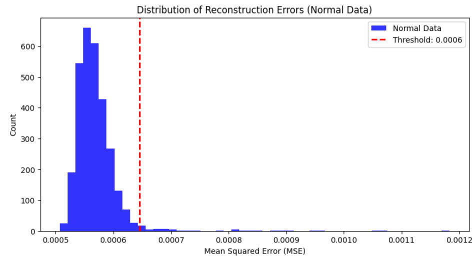

Figure 5: Distribution of Reconstruction Errors (Normal Data)

4.4 Isolation Forest

This section presents the results obtained with the Isolation Forest model applied to the latent representations generated by the autoencoder encoder. The goal is to evaluate how effectively the model can distinguish between normal operating conditions and anomalous fan behavior using only healthy data during training.

4.4.1 Anomaly Score Analysis

Once trained on healthy samples, the Isolation Forest assigns an anomaly score to each test sample using its decision function. Higher anomaly scores indicate a stronger deviation from the learned normal behavior.

The distribution of anomaly scores shows a clear separation between normal and anomalous samples. Normal samples are concentrated around lower anomaly score values, while anomalous samples exhibit significantly higher scores. This separation suggests that the latent features extracted by the encoder preserve discriminative information relevant for anomaly detection.

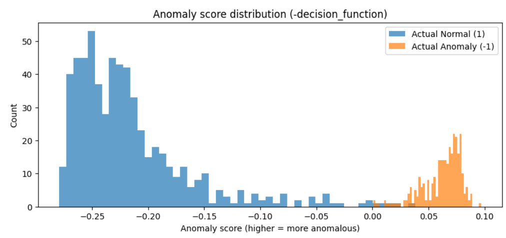

Figure 6: Anomaly score distribution obtained with the Isolation Forest

This behavior confirms that the Isolation Forest is able to model the normal operating region effectively and assign higher anomaly scores to samples corresponding to abnormal conditions.

4.4.2 Classification Performance

To quantitatively assess performance, predictions were compared against the ground-truth labels using a confusion matrix. Out of 602 normal samples in the test set, 597 were correctly classified as normal, while only 5 were incorrectly flagged as anomalous. Importantly, all 300 anomalous samples were correctly detected.

This corresponds to an overall accuracy of approximately 99%, with a False Positive Rate (FPR) of 0.83% and a False Negative Rate (FNR) of 0.00%. The absence of false negatives is particularly relevant in predictive maintenance scenarios, where failing to detect an anomaly may lead to severe operational consequences.

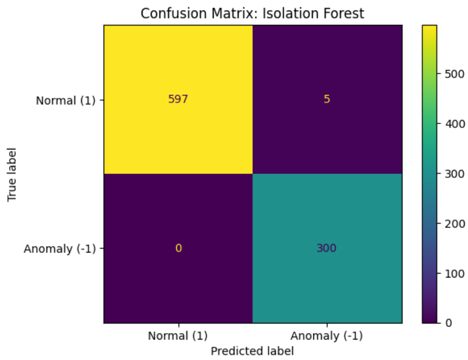

Figure 7: Confusion matrix of the Isolation Forest model

4.4.3 Precision, Recall and F1-Score

To further evaluate the performance of the Isolation Forest model, we analyze the standard classification metrics: precision, recall, and F1-score for both normal and anomalous classes.

As reported in Table 1, the model achieves an overall accuracy of 99% on the test set, which contains both healthy and anomalous samples. For the anomalous class, the model reaches a recall of 1.00, meaning that all anomalous sounds in the test set are correctly detected. This result is particularly important in anomaly detection scenarios, where missing an actual anomaly can have significant practical consequences.

The precision for the anomalous class is 0.98, indicating that only a very small number of normal samples are incorrectly classified as anomalies. This behavior reflects a low false positive rate, which is desirable in predictive maintenance contexts in order to avoid unnecessary alarms or interventions.

For the normal class, the model achieves a precision of 1.00 and a recall of 0.99, confirming that the vast majority of healthy sounds are correctly identified as normal. The resulting F1-scores are close to 1.00 for both classes, demonstrating a well-balanced performance between sensitivity to anomalies and robustness to false alarms.

Table 1: Classification report for the Isolation Forest model on the test set.

Class	Precision	Recall	F1-score	Support
Anomaly (-1)	0.98	1.00	0.99	300
Normal (1)	1.00	0.99	1.00	602
Accuracy			0.99	902
Macro Avg	0.99	1.00	0.99	902
Weighted Avg	0.99	0.99	0.99	902

Overall, these metrics confirm that the Isolation Forest model provides a reliable separation between normal and anomalous operating conditions when applied to the latent representations extracted by the autoencoder.

4.4.4 Error Analysis

A detailed error analysis was conducted to better understand the few misclassifications produced by the model. The five false positives correspond to normal samples that were assigned slightly higher anomaly scores and therefore classified as anomalous. No false negatives were observed.

An inspection of the anomaly scores associated with false positives reveals that these samples lie close to the decision boundary. This suggests that they exhibit acoustic patterns that deviate mildly from the typical normal behavior, potentially representing borderline operating conditions rather than clear anomalies.

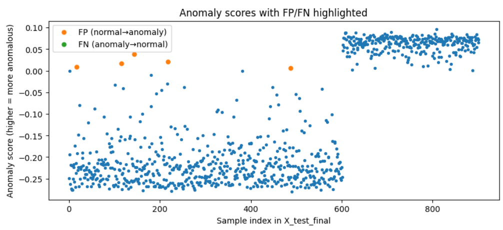

Figure 8: Anomaly scores produced by the Isolation Forest model for all test samples. False positives and false negatives are highlighted.

Overall, the Isolation Forest demonstrates excellent detection capability when combined with autoencoder-based latent representations, achieving near-perfect anomaly detection while maintaining a very low false alarm rate.

4.5 Evaluation of One-Class SVM

This section presents the evaluation of the One-Class Support Vector Machine (OCSVM) model applied to the same test set used for the Isolation Forest, enabling a direct and consistent comparison between both approaches. The evaluation focuses on anomaly score distributions, confusion matrix analysis, and standard classification metrics.

4.5.1 Anomaly Score Distribution

The One-Class SVM outputs a decision function whose values are typically higher for normal samples and lower for anomalous ones. To improve interpretability, an anomaly score is defined as the negative of the decision function, such that higher values correspond to more anomalous behavior:

$$\text{Anomaly score} = -\,\text{decision\_function}(x) \qquad (1)$$

Figure 9 illustrates the anomaly score distribution for normal and anomalous samples. A clear separation between both distributions can be observed. Anomalous samples are concentrated at higher anomaly scores, while normal samples are clustered around lower values. Although a narrow overlap region exists near the decision boundary, the overall separation indicates that the One-Class SVM effectively captures the distinction between normal and anomalous behavior in the latent feature space.

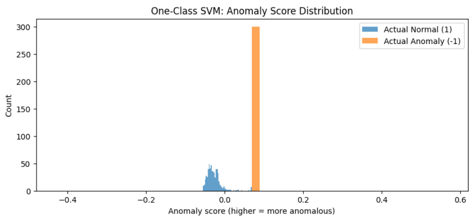

Figure 9: Anomaly score distribution obtained with the One-Class SVM

4.5.2 Confusion Matrix Analysis

The confusion matrix obtained for the One-Class SVM is shown in Figure 10. Out of a total of 902 test samples, all 300 anomalous samples were correctly classified, resulting in zero false negatives. However, 33 normal samples were incorrectly flagged as anomalous, leading to a higher false positive count compared to the Isolation Forest.

This behavior reflects a more conservative detection strategy, where the model prioritizes anomaly detection at the cost of increased false alarms. Such a trade-off may be acceptable in forensic or security-oriented scenarios, where missing anomalous events is more critical than incorrectly flagging normal behavior.

4.5.3 Precision, Recall and F1-Score

Table 2 summarizes the precision, recall, and F1-score obtained by the One-Class SVM.

For the anomalous class, the model achieves a recall of 1.00, indicating that all anomalous samples are successfully detected. The precision of 0.90 reflects the presence of false positives. For the normal class, precision remains perfect, while recall slightly decreases due

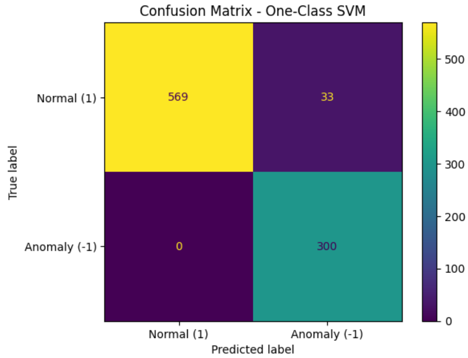

Figure 10: Confusion matrix for the One-Class SVM evaluated on the test set.

to normal samples incorrectly classified as anomalies. The overall accuracy of the model is 0.96.

These results confirm that the One-Class SVM exhibits strong anomaly detection capabilities, particularly in terms of recall for anomalous samples, at the expense of a higher false positive rate compared to the Isolation Forest.

Class	Precision	Recall	F1-score	Support
Anomaly (-1)	0.90	1.00	0.95	300
Normal (1)	1.00	0.95	0.97	602
Accuracy			0.96	902
Macro Avg	0.95	0.97	0.96	902
Weighted Avg	0.97	0.96	0.96	902

Table 2: Classification report for the One-Class SVM evaluated on the test set.

4.5.4 Decision Boundary Visualization and Interpretability

The One-Class SVM is trained in a high-dimensional feature space, where each audio sample is represented by a flattened latent vector extracted from the encoder. In this case, each sample lies in a space of 32 768 dimensions. While this representation is effective for anomaly detection, it makes direct visualization of the learned decision boundary infeasible.

To provide an intuitive interpretation of the model behavior, the feature vectors are projected into a two-dimensional space using Principal Component Analysis (PCA). This projection is applied only for visualization purposes and does not affect the training or evaluation of the One-Class SVM, which is always performed in the original high-dimensional space.

Figure 11 shows the resulting 2D PCA projection of the test samples, distinguishing between normal and anomalous data. As observed, normal samples are compressed into a dense cluster near the origin, while anomalous samples appear separated at a much larger scale along the principal components. This strong scale imbalance leads to a visually distorted representation, where the geometric structure of the decision boundary is difficult to interpret.

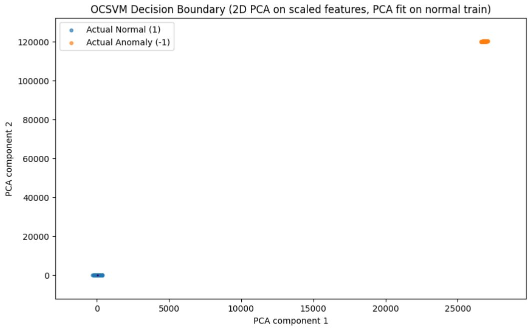

Figure 11: Two-dimensional PCA projection of the One-Class SVM feature space

This behavior is expected and can be explained by two main factors. First, PCA is a variance-preserving technique rather than a class-separating one; it prioritizes directions of maximum variance without considering the decision function of the classifier. Second, projecting an extremely high-dimensional space into only two dimensions inevitably causes a severe loss of geometric information, particularly for non-linear models such as the One-Class SVM with an RBF kernel.

As a result, the apparent boundary in the 2D visualization does not correspond to the true hypersurface learned by the model in the original feature space. Instead, the figure should be interpreted as a qualitative illustration of the data distribution rather than an exact depiction of the classifier decision region.

Importantly, all quantitative evaluation metrics reported for the One-Class SVM, including precision, recall, false positive rate, and false negative rate, are computed using the original high-dimensional features. Therefore, the limitations of the 2D visualization do not affect the validity of the performance results, which remain the primary basis for model comparison and assessment.

4.6 Comparison between Isolation Forest and One-Class SVM

In order to compare the performance of the two anomaly detection approaches under the same conditions, both the Isolation Forest and the One-Class SVM were evaluated using the same test set and the same labeling convention. The comparison focuses on their error behavior, particularly in terms of False Positive Rate (FPR) and False Negative Rate (FNR), which are critical metrics in anomaly detection scenarios.

Table 3 summarizes the results obtained for both models.

Both models achieve a False Negative Rate of zero, indicating that all anomalous samples in the test set are correctly detected. This result shows that, under the selected configuration, neither method fails to identify abnormal behavior, which is a desirable property in safety-critical or security-oriented applications.

Model	**FPR**	**FNR**
Isolation Forest	0.0083 (5 / 602)	0.0000 (0 / 300)
One-Class SVM	0.0548 (33 / 602)	0.0000 (0 / 300)

Table 3: Comparison of False Positive Rate (FPR) and False Negative Rate (FNR) for Isolation Forest and One-Class SVM.

The main difference between the two approaches appears in the False Positive Rate. The Isolation Forest produces a significantly lower number of false alarms, incorrectly classifying only 5 normal samples as anomalous. In contrast, the One-Class SVM misclassifies 33 normal samples, resulting in a higher FPR.

This behavior reflects the different decision mechanisms of both models. While the One-Class SVM constructs a tight boundary around the normal data distribution, making it more sensitive to deviations, the Isolation Forest relies on random partitioning of the feature space, which in this case leads to a more conservative classification of normal samples.

These results highlight a clear trade-off between sensitivity and robustness to false alarms. The implications of this trade-off, as well as the suitability of each approach for different operational contexts, are further discussed in the final conclusions.

5 Conclusion

In this project, we developed and evaluated a hybrid anomaly detection system for predictive maintenance of industrial fans by combining Deep Learning with classical Statistical Machine Learning techniques. This approach addresses the common industrial challenge of highly unbalanced data, where large amounts of healthy recordings are available while faulty examples are scarce.

A Convolutional Autoencoder was trained exclusively on healthy audio data to learn a compact latent representation of normal machine behavior. The low reconstruction error obtained on validation data confirms that the encoder successfully captured the relevant acoustic patterns while reducing the dimensionality of the original spectrograms. This step proved essential, as it enabled the effective application of statistical anomaly detection methods on the extracted latent features.

Two anomaly detection models were evaluated on this latent space. The Isolation Forest achieved excellent performance, detecting all anomalous samples with a very low false positive rate. In contrast, the One-Class SVM also achieved perfect recall on anomalies but produced a higher number of false positives, reflecting a more conservative decision boundary around normal data.

The comparison between both models highlights the trade-off between robustness and sensitivity in anomaly detection. While Isolation Forest appears more suitable for minimizing false alarms, One-Class SVM prioritizes anomaly detection at the cost of increased false positives. Visualization attempts using 2D PCA further illustrate the limitations of representing high-dimensional decision boundaries in low-dimensional spaces.

Overall, the results demonstrate that combining autoencoder-based feature learning with statistical anomaly detection provides an effective and flexible framework for industrial fault detection under limited supervision, making it well suited for real-world predictive maintenance applications.

6 Github Repository

The complete source code, dataset preprocessing scripts, and Jupyter notebooks used in this project are available in the following repository [3]:

`https://github.com/Torres08/dfb-anomaly-detection`

References

[1] Harsh Purohit, Ryo Tanabe, Kenji Ichige, Takashi Endo, Yuki Nikaido, Kaori Suefusa, and Yohei Kawaguchi. Mimii dataset: Sound dataset for malfunctioning industrial machine investigation and inspection. In *Proceedings of the Detection and Classification of Acoustic Scenes and Events 2019 Workshop (DCASE2019)*, pages 209–213, New York, NY, USA, November 2019.

[2] GitHub Repository. Convolutional autoencoder architecture visualization. `https://repository-images.githubusercontent.com/466581223/65cc6a14-02db-4761-bd56-238cf86898df`, 2025. Accessed: 2025-12-27.

[3] Juan Luis Torres Ramos and Rafael Pardo Rueda. Dfb anomaly detection: Predictive maintenance of industrial machinery using deep autoencoders and isolation forest. `https://github.com/Torres08/dfb-anomaly-detection`, 2026. Accessed: 2026-01-19.