applied sciences

MDPI

Article

Innovative Anomaly Detection in PCB Hot-Air Ovens Using Adaptive Temporal Feature Mapping

Chen-Yang Cheng [1], Chuan-Min Chien [1], Tzu-Li Chen [2], Chumpol Yuangyai [3,*] and Pei-ling Kong [1]

[1] Department of Industrial Engineering and Management, National Taipei University of Technology, Taipei 10608, Taiwan; cycheng@ntut.edu.tw (C.-Y.C.); t107379005@ntut.org.tw (C.-M.C.); peang1700@gmail.com (P.-l.K.)
[2] Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu 30013, Taiwan; chentzuli@ie.nthu.edu.tw
[3] Department of Industrial Engineering, School of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand
* Correspondence: chumpol.yu@kmitl.ac.th

Abstract

As automated equipment in PCB manufacturing becomes increasingly reliant on precision hot-air ovens, ensuring operational stability and reducing downtime have become critical challenges. Existing anomaly detection methods, such as Support Vector Machines (SVMs), Deep Neural Networks (DNNs), and Long Short-Term Memory (LSTM) Networks, struggle with high-dimensional dynamic data, leading to inefficiencies and overfitting. To address these issues, this study proposes an innovative anomaly detection system specifically designed for fault diagnosis in PCB hot-air ovens. The motivation is to improve accuracy and efficiency while adapting to dynamic changes in the manufacturing environment. The core innovation lies in the introduction of the Adaptive Temporal Feature Map (ATFM), which dynamically extracts and adjusts key temporal features in real time. By combining ATFM with Bi-Directional Dimensionality Reduction (BDDR) and eXtreme Gradient Boosting (XGBoost), the system effectively handles high-dimensional data and adapts its parameters based on evolving data patterns, significantly enhancing fault detection accuracy and efficiency. The experimental results show a fault prediction accuracy of 99.33%, greatly reducing machine downtime and product defects compared to traditional methods.

Keywords: adaptive temporal feature map; anomaly detection; XGBoost; fault diagnosis; PCB manufacturing

check for updates

Academic Editor: Giorgio De Nunzio

Received: 25 June 2025
Revised: 5 September 2025
Accepted: 10 September 2025
Published: 7 October 2025

Citation: Cheng, C.-Y.; Chien, C.-M.; Chen, T.-L.; Yuangyai, C.; Kong, P.-l. Innovative Anomaly Detection in PCB Hot-Air Ovens Using Adaptive Temporal Feature Mapping. *Appl. Sci.* **2025**, *15*, 10771. https://doi.org/10.3390/app151910771

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The advent of Industry 4.0 has transformed modern manufacturing, aiming to achieve maximum efficiency and performance with minimal resource usage [1]. Originating as a German initiative, Industry 4.0 emphasizes intelligent factories that integrate critical enabling technologies such as the Internet of Things, cyber-physical systems, big data analytics, and artificial intelligence [2]. These technologies enable real-time monitoring, online communication, and rapid decision-making in manufacturing environments [3], thus increasing production efficiency and reducing waste. Within this context, machine learning has emerged as one of the most promising tools for implementing Industry 4.0 practices. In particular, deep learning algorithms have shown superior performance in equipment fault diagnosis compared with traditional approaches, offering higher accuracy in classification tasks [4].

Printed Circuit Board (PCB) manufacturing plays a pivotal role in the electronics sector, serving as the foundation of most electronic systems. PCBs account for the largest share of the global electronic components market due to their ubiquitous use in consumer electronics [5], with Asia producing nearly three-fourths of the world's PCBs. Despite its importance, PCB manufacturing is highly complex [6], involving numerous process variables and sophisticated equipment. Among these, hot-air ovens are critical to production but prone to frequent failures, such as motor faults, heater disconnection, or bearing defects. Such anomalies can cause product damage, equipment downtime, and significant economic losses. Traditional fault detection approaches, which rely on manual monitoring of oven signals, are inefficient, labor-intensive, and prone to human error.

Recent advances in data-driven anomaly detection provide new opportunities to address these challenges. Automated systems employing machine learning and data mining techniques can process large-scale production data more effectively, enabling early fault detection and preventive maintenance. However, existing solutions still face limitations in handling high-dimensional, imbalanced, and evolving production data.

To overcome these issues, this study proposes an innovative framework that integrates Adaptive Temporal Feature Mapping (ATFM) with Bi-Directional Dimensionality Reduction (BDDR) and eXtreme Gradient Boosting (XGBoost). The ATFM dynamically extracts and adjusts temporal features in real time, while the combined use of BDDR and XGBoost enhances classification accuracy and robustness. By automating anomaly detection in PCB hot-air ovens, the proposed system not only identifies existing faults promptly but also predicts potential failures, thereby minimizing production downtime and improving manufacturing efficiency.

The remainder of this research is structured as follows: Section 2 provides a literature review of relevant research, including the theoretical concepts in anomaly detection based on machine learning. Section 3 explains the architecture of BDDR-XGBoost for the preventive diagnosis of oven machines. Section 4 presents the experiments and prediction results, followed by conclusions and suggestions for future work in Section 5.

2. Literature Review

With the increasing complexity of manufacturing processes and machines, the need for effective and efficient techniques to monitor real-time conditions has grown significantly [7]. Anomaly detection, a crucial data analysis task, aims to identify unusual or abnormal data from a given dataset [8]. It has been widely studied in computer vision, statistics, and machine learning [9] and applied across various industries, including the chemical industry [10], food industry [11], medical and public health, fraud detection, intrusion detection, image processing, sensor networks, and robotics [12]. In industrial contexts, anomaly detection systems play a crucial role in identifying machine malfunctions and underperforming tools that may otherwise be undetected for a long time using conventional approaches, providing real-time alerts that reduce production losses and prevent equipment failures [13,14].

To address anomaly detection in manufacturing, numerous machine learning algorithms have been adopted, such as Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), Decision Trees, XGBoost, logistic regression [15], Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) networks [16]. Comparative studies highlight their respective advantages and drawbacks. For example, SVMs can handle high-dimensional data but require extensive memory and computational resources. Neural networks often achieve strong performance but involve long training times for large architectures. Decision trees are relatively fast but tend to produce over-complex models that lack generalization [17]. These limitations hinder

their application in large-scale industrial environments such as Printed Circuit Board (PCB) manufacturing, where the data are high-dimensional, imbalanced, and time-sensitive.

Recent studies have demonstrated the advantages of XGBoost for anomaly detection. An XGBoost–DNN model outperformed Naive Bayes, SVM, and Logistic Regression in intrusion detection tasks. In the field of offshore wind turbine fault detection, XGBoost achieved higher predictive accuracy and computational efficiency compared with LSTM [18]. XGBoost has proven effective in addressing data-overfitting problems, handling time-series data [18], and providing execution speed and memory efficiency [19]. Its use of first- and second-order gradient approximations also helps stabilize the learning process and minimize loss functions [20].

Accordingly, XGBoost was selected in this study as the core classifier for anomaly detection in PCB manufacturing. Its computational efficiency, robustness, and scalability make it particularly suitable for tackling the challenges of PCB production data. To this end, we propose a BDDR-XGBoost framework that integrates statistical and symbolic dimensionality reduction techniques with XGBoost, enabling efficient and accurate anomaly detection in industrial processes.

3. ATFM-BDDR-XGBoost Framework

This section introduces the proposed ATFM-BDDR-XGBoost model for feature extraction and fault prediction in the PCB industry. The model consists of five key components: data preprocessing, data transformation, handling of imbalanced datasets, prediction modeling, and evaluation metrics.

3.1. Data Preprocessing

This study adopts a comprehensive data preprocessing approach to address fault identification in the Printed Circuit Board (PCB) industry. The dataset was collected between November 2020 and September 2021, covering multiple machine attributes as listed in Table 1. In total, the dataset comprises 210 production cycles, with each cycle consisting of 198 data points, resulting in 41,580 individual records.

Table 1. Data analysis related attributes.

Attributes	Data Types
Time	Time (hh:mm:ss)
Current temperature (°C)	Numeric
Electricity consumption by windmill motors (A)	Numeric
Electricity consumption by heaters (A)	Numeric

The machine operating status is categorized into five classes: normal production, heater disconnection, wire cutting angle of the heater, abnormal motor noise, and motor insulation failure. Among these, normal production accounts for the majority of the data, while the four fault categories provide sufficient abnormal cases to support supervised anomaly detection (Figure 1). This distribution ensures that both normal and faulty conditions are adequately represented for model training and evaluation. Each production cycle comprises three distinct phases: preheating (20 groups), steady-state (40 groups), and cooling (6 groups), as shown in Figure 2. During these phases, key process variables—including temperature (≈approximately 40–150 °C), electric current from windmill motors (≈approximately 2.0–2.4 A), electric current from heaters (≈approximately 0.05–16 A), and processing time—are continuously recorded. These ranges reflect realistic operating conditions of the oven system.

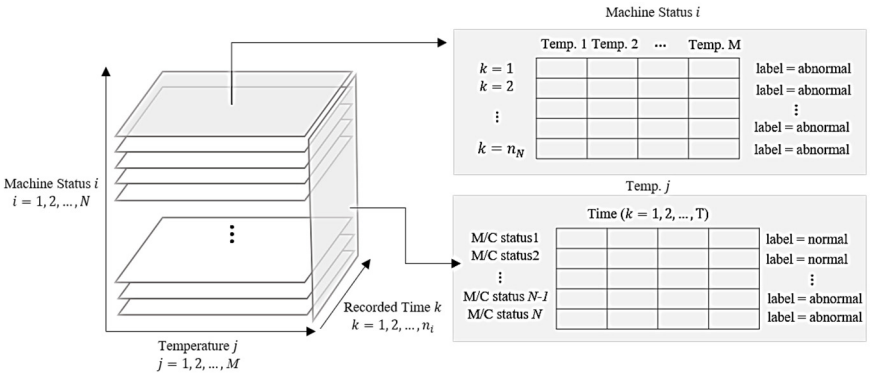


Figure 1. Illustration of raw data transposed into input formats.

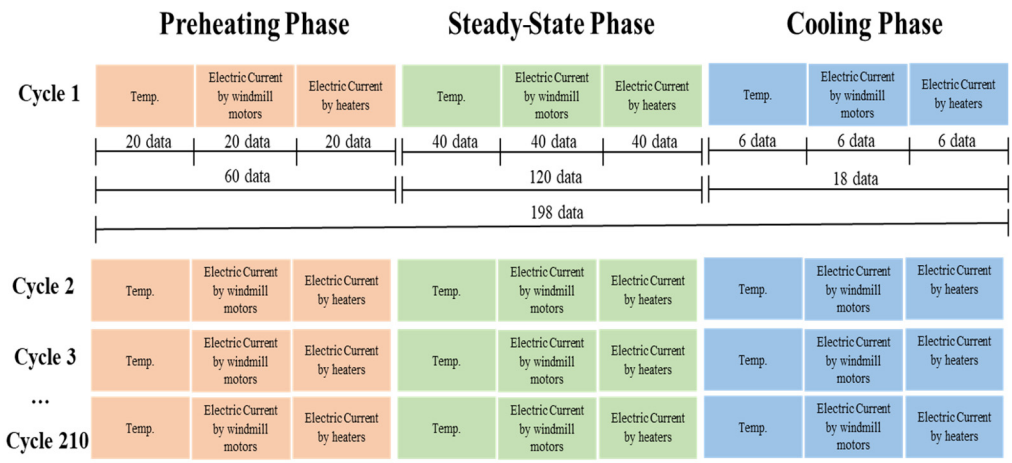


Figure 2. Selecting appropriate operating data.

To prepare the dataset for analysis, missing values were removed, and the remaining records were averaged within each phase group. Before applying dimensionality reduction methods such as Symbolic Aggregate Approximation (SAX), each time series was standardized using z-normalization, ensuring that the mean was 0 and the standard deviation was 1. While this transformation does not guarantee a perfectly Gaussian distribution, it sufficiently approximates the assumption underlying SAX, as supported in prior studies. This step helps stabilize variability across different signals and enables a consistent symbolic representation. The grouped data were then rescaled and reformatted into a standardized structure for anomaly detection modeling. This preprocessing ensures that the dataset is both representative of actual production conditions and suitable for machine learning analysis.

3.2. Adaptive Temporal Feature Mapping (ATFM) Integration

To enhance the predictive accuracy and adaptability of the XGBoost model, this research integrates an Adaptive Temporal Feature Map (ATFM) into the anomaly detection framework. The ATFM captures the evolving characteristics of the data within specific time windows, enabling the model to adjust to time-dependent patterns dynamically. This is particularly effective in detecting anomalies in PCB manufacturing, where production cycles and machine statuses change over time. The ATFM enhances conventional dimensionality reduction techniques such as Principal Component Analysis (PCA) and Symbolic Aggregate Approximation (SAX) by introducing a temporal dynamic aspect. While PCA and SAX reduce data dimensions statically by transforming columns and rows, respectively, ATFM focuses on capturing temporal changes in data within defined time windows, thereby enabling dynamic analysis of machine operations.

Step 1: Define Time Windows

The data are divided into time windows to track dynamic changes. Each time window T_w consists of a set of consecutive data points:

$$T_w = \{t_1, t_1, \ldots t_n\} \tag{1}$$

For this research, time windows are defined for each of the three production phases (heating, steady-state, and cooling) to ensure that the features extracted reflect the phase-specific patterns.

Step 2: Extract Features

Within each time window tit_iti, the following key features are extracted:

- Mean (μ_{t_i}) The average value of the data points in the window.
- Standard deviation (σ_{t_i}) The dispersion of the data within the window indicates variability.
- Rate of change (Δ_{t_i}): The change in data between consecutive time windows.

The formulas used to calculate these features are as follows:

$$\mu_{t_i} = \frac{1}{n}\sum_{j-1}^{n} x_j \tag{2}$$

$$\sigma_{t_i} = \sqrt{\frac{1}{n}\sum_{j-1}^{n}\left(x_j - \mu_{t_i}\right)^2} \tag{3}$$

$$\Delta_{t_i} = \frac{x_{t_i} - x_{t_{i-1}}}{t_i - t_{i-1}} \tag{4}$$

where x_j represents data points within the time window t_i, and n is the number of points in the window.

Step 3: Generate the Temporal Feature Map (TFM)

Once these features are computed for each time window, they form the Temporal Feature Map (TFM):

$$\mathrm{F}_T = \left\{F_{t_1}, F_{t_2}, \ldots F_{t_n,}\right\} \tag{5}$$

where each F_T represents the set of features $\{\mu_{t_i}, \sigma_{t_i}, \Delta_{t_i}\}$ for the corresponding time window. This TFM offers a comprehensive view of the temporal behavior of the data and is particularly useful for monitoring machines during various production phases.

Step 4: Integrate with XGBoost

The ATFM is then integrated into the XGBoost model to dynamically adjust its parameters according to the evolving data. Unlike traditional XGBoost models that use static features, the inclusion of ATFM allows the model to be adaptive, refining its predictions over time by adjusting to the real-time state of the machine. The optimal model parameters θ_{t_i} for each time window, the parameters are determined by minimizing the loss function:

$$\theta_{t_i} = \underset{\theta}{argmin}\, L(F_{t_i}, \theta) \tag{6}$$

where $\mathrm{F}_T = \{\mu_{t_i}, \sigma_{t_i}, \Delta_{t_i}\}$ represents the dynamic features for the time window t_i, and L is the loss function to be minimized.

Step 5: Visualize Temporal Feature Map (TFM)

To better understand the evolving characteristics of the data, the ATFM can be visualized by plotting the dynamic features (mean, standard deviation, and rate of change) over

time. This visualization provides insights into how the data behaves in various production phases, enabling the detection of anomalies more effectively.

3.3. Imbalance Dataset

In real-world domains, several oversampling techniques have proven effective. Among these techniques, the Synthetic Minority Oversampling Technique (SMOTE) has gained significant popularity as it interpolates existing samples to generate new instances [7]. This method's simplicity and computational efficiency have made it a popular choice among researchers, and studies have shown that selective oversampling using SMOTE leads to improved classification performance [7–9]. In this paper, we address the issue of label imbalance in the dataset and adopt the SMOTE algorithm to duplicate the minority class randomly until it is equivalent to the majority class, thereby balancing the dataset before training the model. Figure 3 illustrates the process of generating synthetic data points using the SMOTE algorithm. This is achieved by linearly interpolating a randomly selected minority observation and one of its neighboring minority observations. The SMOTE calculation in our research generates new samples (x_{new}) as shown in Equation (7), where x_i is a selected minority sample, x_j is a minority sample, and w is a random weight between 0 and 1 [7–10].

$$x_{new} = x_i + w \times (x_j - x_i) \tag{7}$$

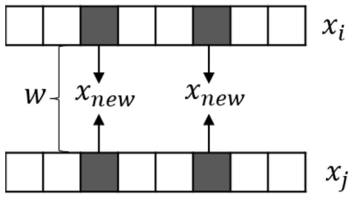


Figure 3. Synthetic data generation in the SMOTE algorithm.

3.4. Prediction Model

The XGBoost algorithm utilizes a tree ensemble model, comprising a set of classification and regression trees. XGBoost has been proven to surpass the limits of computing power for boosted tree algorithms, making it a popular choice among researchers and data scientists [13]. The core concept of the boosting algorithm is to integrate multiple classifiers and convert weak learners into strong learners. During each iteration of gradient boosting, the residual is used to correct the previous tree, optimizing the specified loss function [14]. XGBoost is distinct from traditional decision tree boosting algorithms, such as GBDT (Gradient Boosting Decision Tree) [15], and it is designed to be flexible and computationally efficient [16], with the ability to perform time-series forecasting.

This research utilizes the XGBoost algorithm to predict failures in critical components of oven machine tool systems, specifically targeting windmill motors and heaters. Training of the algorithm was conducted on Python 3.6.8, incorporating various scientific computing libraries for data preprocessing and predictive modeling. The experiments were carried out on a Windows 10 system equipped with an Intel (R) Core (TM) i7-9750H CPU, operating at 2.60 GHz and 2.59 GHz.

The dataset was split into 80% for training and 20% for validation to calibrate the model. This study utilized Adaptive Temporal Feature Mapping (ATFM), which dynamically extracts key features such as mean, standard deviation, and rate of change from the data across time windows. These temporal features were then used to adaptively adjust key hyperparameters, including gamma, learning_rate, and n_estimators, throughout the training process. Unlike traditional methods such as Grid Search Cross-Validation, this

adaptive approach enables the XGBoost model to continuously fine-tune its parameters based on evolving data characteristics. The model was trained by sequentially adding trees, where each tree responded to the residual errors and the temporal dynamics captured by ATFM, resulting in a more flexible and accurate composite model for fault prediction.

3.5. Evaluation Metrics

In this study, the proposed model's performance was evaluated using the accuracy metric, which calculates the number of correctly predicted data points out of all the data points. For multiclass classification, accuracy can also be calculated in terms of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) using Equation (8).

$$Accuracy = \frac{\sum TP + \sum TN}{\sum TP + \sum TN + \sum FP + \sum FN} \quad (8)$$

To evaluate the model's performance and visualize the relationship between multiclass classification, a 5×5 confusion matrix was employed in this research.

The dynamic and real-time adjustment of temporal features in the proposed ATFM is achieved through continuous updates of the temporal windows and the feature extraction process. As production data stream in, each phase (preheating, steady-state, and cooling) is rescaled into fixed-length groups, ensuring that cycles of varying lengths are standardized. Within each group, statistical features such as the mean, standard deviation, and rate of change are recalculated in real time, thereby adapting the feature representation to the latest machine conditions. These evolving features are compiled into a Temporal Feature Map (TFM), which is directly integrated into the XGBoost classifier. Because XGBoost is computationally efficient, the model can promptly incorporate the updated feature map, enabling the system to dynamically adapt to temporal variations and detect anomalies as soon as they occur. This integration ensures that the anomaly detection framework does not rely on static feature sets, but rather adjusts continuously and in real time to the operational state of the hot-air oven machine.

4. Experimental Results

This section is organized into three integral parts. Initially, it examines the impact of various conditions on the predictive accuracy of the proposed model, providing insights into its operational adaptability. Following this, a comparative analysis is conducted to evaluate the performance of the XGBoost algorithm in relation to other contemporary machine learning techniques. This comparison is vital to contextualize the model's efficacy within the broader machine learning spectrum. The final part of this section focuses on evaluating the model's performance, both with and without the application of dimensionality reduction techniques. This evaluation is key to determining the effectiveness of these techniques in enhancing the model's predictive capabilities and overall efficiency.

4.1. Dataset Details and Results Comparison

The dataset used in this study covers production data from November 2020 to September 2021. All machine attributes, such as temperature and electric current, were recorded. These data are divided into three production phases: preheating, steady-state, and cooling, with varying data sizes for each phase. This study describes experiments on data rescaling based on data size, including small size (99 data points), medium size (198 data points), and large size (258 data points). Principal Component Analysis (PCA) was applied to reduce data dimensionality, and the impact of different numbers of principal components on prediction results was compared. The results showed an accuracy of 94.00% when the number of components was automatically set, and 97.33% and 86.66% when manually set

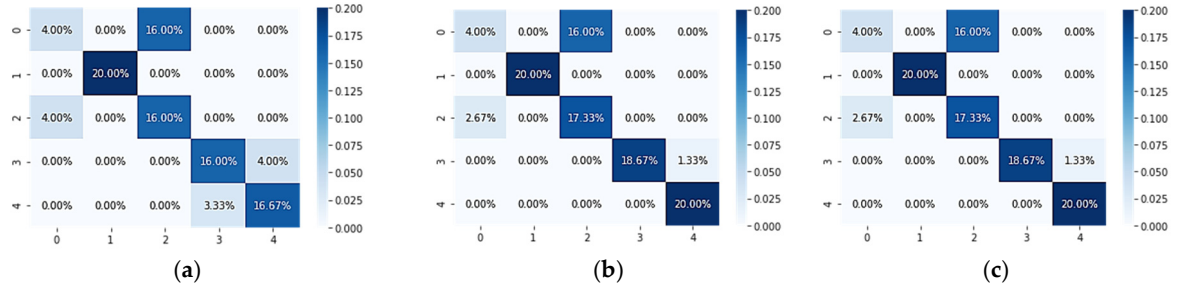
to 20 and 70 principal components, respectively. This study also conducted experiments using the Symbolic Aggregate Approximation (SAX) technique and compared the effect of varying the number of symbols on the prediction model. The best prediction accuracy was 64.29% when numerical data were converted to 25 and 30 different symbols. The prediction results of various model algorithms, including XGBoost, ANN, and LSTM, were compared. The results indicated that the XGBoost model outperformed the ANN and LSTM models with an average accuracy of 80%. This study presents a detailed comparison of prediction results obtained using different dimensionality reduction techniques in conjunction with the XGBoost model. The results show that the BDDR-XGBoost model, which combines PCA and SAX techniques, achieves the best prediction performance, with an accuracy of 99.33%.

4.2. Comparison of Prediction Results with the Data Size

This section outlines the results of various experiments designed to evaluate the proposed model's performance under different operational conditions. Data were categorized into three scales—small (99 data points), medium (198 data points), and large (258 data points)—corresponding to the preheating, steady-state, and cooling phases. A systematic comparison of the prediction performance for each dataset is presented in Table 2. Furthermore, to facilitate a deeper understanding of these results, confusion matrices for each dataset size are depicted in Figure 4a–c. It was observed that the medium-sized dataset, comprising 198 rescaled data points in a single cycle, exhibited enhanced performance compared to the smaller and larger datasets.

Table 2. Performance comparison between different amounts of rescaling data.

Size of Datasets	Number of Each Phase			Number of Data with 1 Cycle	Accuracy (%)
	Preheating	Steady State	Cooling		
Small	10	20	3	99	72.67%
Medium	20	40	6	198	80.00%
Large	30	50	6	258	70.00%



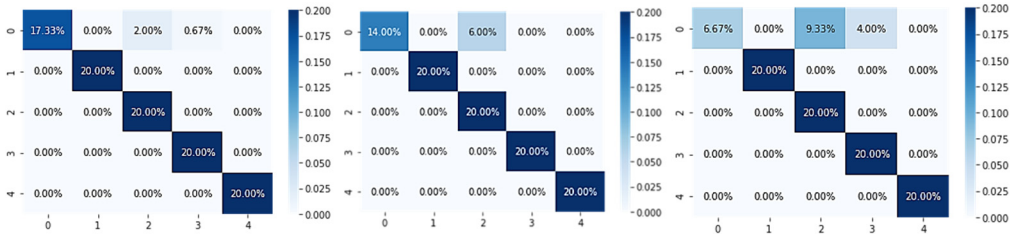
(**a**) (**b**) (**c**)

Figure 4. Confusion matrices for datasets of different rescaling sizes: (**a**) small (99 data points); (**b**) medium (198 data points); (**c**) large (258 data points).

In addition to size variation, Principal Component Analysis (PCA) was employed to reduce data dimensionality. The number of principal components was determined based on achieving a variance proportion of 0.95, thereby accounting for 95% of the original dataset's variance. A comparative study of the model's accuracy with varying numbers of principal components was undertaken. The findings, as tabulated in Table 3, reveal that model accuracy was 94.00% with automatically set principal components, it increased to 97.33% with 20 components, and it decreased to 86.66% with 70 components. The associated confusion matrices, illustrating these variations, are provided in Figure 5a–c. Based on these findings, 20 was identified as the optimal number of principal components.

Table 3. Performance comparison between different numbers of principal components.

Size of Principal Components	Number of Components	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Small	20	97.33	97.53	97.33	97.29
Medium	28	94	95.38	94	93.86
Large	70	86.66	90.3	86.67	84.4



(**a**) (**b**) (**c**)

Figure 5. Confusion matrices for experiments with different numbers of principal components: (**a**) 20 components; (**b**) 28 components; (**c**) 70 components.

4.3. Comparison of Prediction Results with the Different Models

This section details the empirical findings of our investigation, which was focused on the fault prediction in hot-air oven datasets utilizing a range of algorithmic models. The core of this comparative study was an analysis of the performance metrics of three distinct models: ANNs, LSTM networks, and XGBoost. To provide a comprehensive evaluation, a series of 15 distinct testing iterations were executed for each algorithm. The results from these experimental assessments are systematically encapsulated in Table 4. This table offers an in-depth comparison of the predictive accuracies achieved by each algorithm under study. A salient observation from our analysis is the exceptional performance of the XGBoost model, which consistently recorded an average accuracy of 80%. This marked a significant improvement over the results obtained from the ANN and LSTM models. In order to present a more tangible and interpretable view of each model's predictive efficacy, confusion matrices were employed. These matrices are elucidated in Figure 6a–c, corresponding to the ANN, LSTM, and XGBoost models, respectively. The confusion matrices serve as an illustrative tool, providing a clear visual representation of the predictive strengths and weaknesses of each model. Based on the comprehensive data gathered and the ensuing analysis, it can be confidently posited that the XGBoost model demonstrates a superior capability in fault prediction across the various datasets evaluated in this study.

Table 4. Comparison of prediction results with the different models.

Model	Accuracy (%)
ANN	20.00%
LSTM	40.00%
XGBoost	80.00%
SAX-XGBoost	69.33%
PCA-XGBoost	98.67%
ATFM-BDDR-XGBoost	99.33%

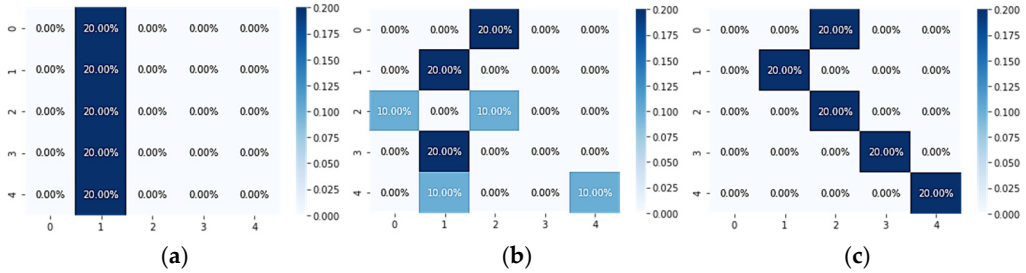


Figure 6. Confusion matrices for different model algorithms: (**a**) ANN model; (**b**) LSTM model; (**c**) XGBoost model.

4.4. Comparison of Prediction Results with the Different Conditions

In the empirical analysis, the integration of the XGBoost model with the Symbolic Aggregate Approximation (SAX) technique was explored. This approach involved transforming numerical data into 25 distinct symbolic representations for model training purposes. The experimental outcomes yielded an average classification accuracy of 69.33%, with a misclassification rate of approximately 30.67% (46 out of 150 instances). The corresponding confusion matrix, delineated in Figure 7a, elucidates the intricate multiclass relationships inherent in the SAX-XGBoost model.

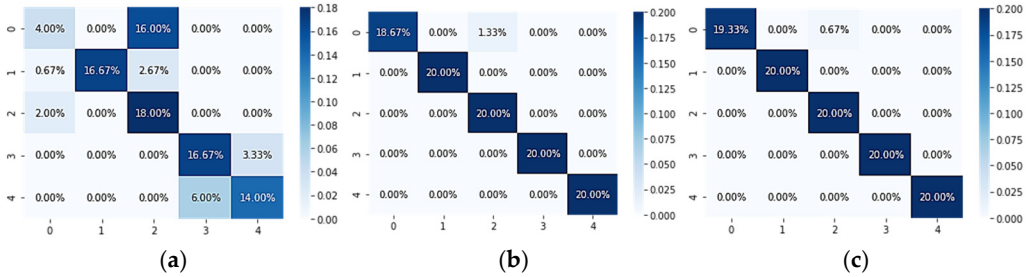


Figure 7. Confusion matrices for combined methods: (**a**) SAX only; (**b**) PCA only; (**c**) PCA + SAX.

Subsequently, an experiment incorporating Principal Component Analysis (PCA) with the XGBoost model was conducted. Here, the PCA's dimensional reduction was deliberately confined to 20 principal components. This PCA-XGBoost amalgamation resulted in a significant enhancement of predictive accuracy, achieving an average of 98.67% with only two instances of misclassification, notably in predicting the normal production condition for the heater's wire cutting angle. The associated confusion matrix is depicted in Figure 7b, highlighting the PCA's pivotal role in mitigating overfitting issues and enhancing prediction accuracy.

Moreover, this study proposed a Bi-Directional Dimensionality Reduction (BDDR) approach combined with the XGBoost model, which demonstrated the highest predictive efficacy. The BDDR-XGBoost model achieved an exceptional average accuracy of 99.33%, with only a single instance of misclassification, specifically in the context of the heater's wire cutting angle during normal production. This model's predictive capability is anticipated to refine fault detection accuracy significantly. This methodology has been employed in the comparative performance analysis of various fault classification algorithms in hot-air oven machines. The performance analysis is systematically presented in a matrix format, comprising five rows and five columns, which depict scenarios of one normal production prediction and four distinct types of abnormal production predictions. The confusion matrix for the BDDR-XGBoost model is exhibited in Figure 7c.

4.5. Comparison of Prediction Results with the Different Methods

In anomaly detection, AUC (Area Under the ROC Curve) is a widely recognized evaluation metric that offers a comprehensive measure of a model's performance. Unlike accuracy, which can be misleading in imbalanced datasets, AUC evaluates the trade-off between the true positive rate (sensitivity) and the false positive rate (1-specificity) across all possible classification thresholds. This makes AUC particularly valuable for assessing models in preventive diagnostics, where the accurate identification of anomalies is crucial. In this study, we complement traditional evaluation metrics, such as accuracy and confusion matrix, with AUC to provide a more holistic assessment of our models' effectiveness in fault detection, as shown in Table 5. These results highlight the superior performance of the proposed model and XGBoost-PCA methods, both achieving near-perfect AUC values close to 1. This indicates that these methods have high predictive accuracy and are highly effective in identifying and diagnosing faults in PCB hot-air ovens. Incorporating AUC alongside accuracy and confusion matrix metrics provides a more robust and comprehensive evaluation of the models' performance in preventive diagnostics.

Table 5. Comparison of AUC results with the different models.

Method	True Positive Rate (TPR)	False Positive Rate (FPR)	AUC	Accuracy (%)
XGBoost	0.8	0.2	0.8	80
SAX-XGBoost	0.69	0.31	0.69	69.33
XGBoost-PCA	0.99	0.01	0.99	98.67
Proposed Model	0.9933	0.0067	0.99	99.33

4.6. Statistical Analysis

In this study, a paired t-test was conducted to determine whether the XGBoost model significantly outperforms the ANN and LSTM models in terms of classification accuracy. Accuracy scores from multiple runs or cross-validation folds were collected for each model, resulting in distributions of accuracy data. Specifically, the accuracy scores from ten runs for each model were as follows: XGBoost [0.82, 0.83, 0.81, 0.80, 0.84, 0.79, 0.85, 0.82, 0.83, 0.81], ANN [0.20, 0.21, 0.19, 0.20, 0.22, 0.18, 0.23, 0.20, 0.21, 0.19], and LSTM [0.40, 0.41, 0.39, 0.38, 0.42, 0.37, 0.43, 0.40, 0.41, 0.39].

The hypothesis test was structured with the null hypothesis (H0) asserting no significant difference in accuracy between the models and the alternative hypothesis (H1) asserting a significant difference. The results of the paired t-test showed a t-statistic of 289.08 and a p-value of 3.61×10^{-19} for the comparison between XGBoost and ANN. For the comparison between XGBoost and LSTM, the t-statistic was 2.07×10^{16}, with a p-value of 7.23×10^{-144}. As both p-values are far below the significance threshold of 0.05, the null hypothesis was rejected, indicating a significant difference in accuracy between the models.

These results clearly demonstrate that the XGBoost model's accuracy is significantly higher than that of both the ANN and LSTM models. This confirms the XGBoost model's superiority in classification tasks, particularly in fault prediction and anomaly detection, contributing to enhanced reliability and efficiency in PCB hot-air oven operations.

4.7. Discussion

The results of this study highlight the superior performance of the ATFM-BDDR-XGBoost framework in detecting anomalies in PCB hot-air ovens. By dynamically extracting and adjusting temporal features, ATFM enabled the system to adapt to evolving data patterns, a capability that traditional models such as ANN and LSTM struggled to achieve. The integration of BDDR further enhanced the framework by combining PCA and SAX,

effectively managing high-dimensional sensor data and improving classification robustness. Together with SMOTE for data balancing, the model achieved a classification accuracy of 99.33%, substantially outperforming the baseline approaches. Beyond predictive accuracy, these improvements indicate that the system can reduce downtime and product defects, while also supporting earlier fault identification, thereby enhancing decision-making and preventive maintenance.

A further analysis of false positives (FP) and false negatives (FN) provides additional insights into the model's reliability. The results from the confusion matrices indicate that the ANN produced a high number of false negatives, meaning that many faulty cases were incorrectly classified as normal, which is particularly dangerous in industrial applications. LSTM reduced false negatives compared to the ANN, but still exhibited poor sensitivity to minority fault categories. In contrast, XGBoost and the proposed ATFM-BDDR-XGBoost framework achieved much lower false negative rates, ensuring that genuine faults were rarely missed. Although this came at the cost of slightly higher false positives (i.e., false alarms), in manufacturing environments, it is generally more acceptable to deal with occasional false alarms than to risk undetected failures. This balance between FP and FN demonstrates that the proposed framework not only improves accuracy but also enhances operational safety by prioritizing fault detection over false alarms.

Several additional aspects warrant discussion. First, the interpretability of extracted features remains a challenge. In this study, statistical features such as mean, standard deviation, and rate of change were chosen based on engineering knowledge; however, a systematic feature importance analysis is needed to validate these assumptions. Future work will incorporate explainable AI tools such as SHAP or LIME to strengthen transparency and link feature behavior with specific fault mechanisms.

Second, although the system is designed for real-time deployment, validation was conducted using historical data in an offline setting. This limitation prevents conclusions about latency and computational feasibility under streaming conditions. Although the framework currently runs efficiently on a standard workstation, additional experiments are required to benchmark its performance in resource-constrained industrial environments. Real-time testing and profiling will therefore be prioritized in future work.

Third, scalability and generalizability require attention. While the approach is based on generic time-series sensor data and thus holds potential for extension to motors, compressors, CNC machines, and chemical processes, its robustness across different operating environments remains untested. Scaling to thousands of sensors on a production line could introduce bottlenecks in data transfer, storage, and retraining. Distributed data architectures and parallelized training strategies will be needed to address these issues.

Finally, the limitations of the dataset must be acknowledged. The absence of extremely rare catastrophic failures may lead to an underestimation of such events. Although SMOTE alleviated class imbalance among observed categories, it cannot substitute for missing fault types. Future directions include synthetic data generation, simulated rare events, and transfer learning to enhance detection capability in underrepresented cases. Moreover, benchmarking against modern temporal modeling strategies such as Temporal Convolutional Networks and Transformer-based architectures will help determine the competitiveness of ATFM-BDDR-XGBoost in comparison to emerging techniques.

5. Conclusions

This study presented an integrated anomaly detection framework for PCB hot-air ovens that combines Adaptive Temporal Feature Mapping (ATFM), Bi-Directional Dimensionality Reduction (BDDR), and XGBoost. The main contribution lies in demonstrating that established techniques can be effectively combined and applied to real industrial data,

with ATFM enabling dynamic feature extraction that enhances adaptability to evolving machine conditions. The framework achieved a fault classification accuracy of 99.33%, highlighting its potential to reduce downtime, prevent product defects, and improve the overall reliability of PCB manufacturing processes. Notably, the system also provides actionable value by supporting preventive maintenance and early decision-making, which are critical in smart manufacturing environments. While further validation on real-time production streams and across other types of industrial equipment is required, the results establish a solid foundation for broader application. Future extensions to other machinery and benchmarking against more modern temporal models will help confirm the framework's scalability and competitiveness. In conclusion, ATFM-BDDR-XGBoost offers a practical and interpretable approach to anomaly detection that bridges academic techniques with industrial needs, contributing to safer operations, higher product quality, and the advancement of intelligent manufacturing systems.

Author Contributions: All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by C.-Y.C., C.Y., P.-l.K., C.-M.C. and T.-L.C. The first draft of the manuscript was written by C.-Y.C. and all authors commented on previous versions of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financially supported by King Mongkut's Institute of Technology Ladkrabang under grant numbers 2566-02-01-003 and NTUT-KMITL-112-02.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Jamwal, A.; Agrawal, R.; Sharma, M.; Giallanza, A. Industry 4.0 technologies for manufacturing sustainability: A systematic review and future research directions. *Appl. Sci.* **2021**, *11*, 5725. [CrossRef]
2. Zhong, R.Y.; Xu, X.; Klotz, E.; Newman, S.T. Intelligent manufacturing in the context of industry 4.0: A review. *Engineering* **2017**, *3*, 616–630. [CrossRef]
3. Kim, J.H. A review of cyber-physical system research relevant to the emerging IT trends: Industry 4.0, IoT, big data, and cloud computing. *J. Ind. Integr. Manag.* **2017**, *2*, 1750011. [CrossRef]
4. Manikandan, S.; Duraivelu, K. Fault diagnosis of various rotating equipment using machine learning approaches—A review. *Proc. Inst. Mech. Eng. Part E J. Process Mech. Eng.* **2020**, *235*, 629–642. [CrossRef]
5. Zhou, Y.; Qiu, K. A new technology for recycling materials from waste printed circuit boards. *J. Hazard. Mater.* **2010**, *175*, 823–828. [CrossRef] [PubMed]
6. LaDou, J. Printed circuit board industry. *Int. J. Hyg. Environ. Health* **2006**, *209*, 211–219. [CrossRef] [PubMed]
7. Douzas, G.; Bacao, F.; Last, F. Improving imbalanced learning through a heuristic oversampling method based on k-means and SMOTE. *Inf. Sci.* **2018**, *465*, 1–20. [CrossRef]
8. Ahmed, M.; Mahmood, A.N.; Hu, J. A survey of network anomaly detection techniques. *J. Netw. Comput. Appl.* **2016**, *60*, 19–31. [CrossRef]
9. Maldonado, S.; López, J.; Vairetti, C. An alternative SMOTE oversampling strategy for high-dimensional datasets. *Appl. Soft Comput.* **2019**, *76*, 380–389. [CrossRef]
10. Skryjomski, P.; Krawczyk, B. Influence of minority class instance types on SMOTE imbalanced data oversampling. In *First International Workshop on Learning with Imbalanced Domains: Theory and Applications*; PMLR: New York, NY, USA, 2017; pp. 7–21.
11. Luengo, J.; Fernández, A.; García, S.; Herrera, F. Addressing data complexity for imbalanced data sets: Analysis of SMOTE-based oversampling and evolutionary undersampling. *Soft Comput.* **2010**, *15*, 1909–1936. [CrossRef]
12. Chen, T.; Guestrin, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
13. Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. *Classification and Regression Trees*; Routledge: Oxfordshire, UK, 2017.
14. Kan, C.; Cheng, C.; Yang, H. Heterogeneous recurrence monitoring of dynamic transients in ultraprecision machining processes. *J. Manuf. Syst.* **2016**, *41*, 178–187. [CrossRef]

15. Ogunleye, A.; Wang, Q.-G. XGBoost model for chronic kidney disease diagnosis. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2019**, *17*, 2131–2140. [CrossRef] [PubMed]
16. Qian, B.; Lu, S. Detection of mobile network abnormality using deep learning models on massive network measurement data. *Comput. Netw.* **2021**, *201*, 108571. [CrossRef]
17. Omar, S.; Ngadi, A.; Jebur, H.H. Machine learning techniques for anomaly detection: An overview. *Int. J. Comput. Appl.* **2013**, *79*, 33–41. [CrossRef]
18. Chen, L.; Gao, X.; Li, X. Using the motor power and XGBoost to diagnose working states of a sucker rod pump. *J. Pet. Sci. Eng.* **2021**, *199*, 108329. [CrossRef]
19. Trizoglou, P.; Liu, X.; Lin, Z. Fault detection by an ensemble framework of Extreme Gradient Boosting (XGBoost) in the operation of offshore wind turbines. *Renew. Energy* **2021**, *179*, 945–962. [CrossRef]
20. Dhaliwal, S.S.; Nahid, A.-A.; Abbas, R. Effective intrusion detection system using XGBoost. *Information* **2018**, *9*, 149. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.