Journal of Industrial Engineering and Applied Science
Journal Home: http://jieas.suaspress.org/ | CODEN: JIEAAE
Vol. 3, No. 6, 2025 | ISSN 3005-6071 (Print) | ISSN 3005-608X (Online)

Predictive Maintenance of Semiconductor Equipment Using Stacking Classifiers and Explainable AI: A Synthetic Data Approach for Fault Detection and Severity Classification

YIN, Min [1*]

[1] University of California-Berkeley, USA

**** YIN, Min is the corresponding author, E-mail: gmiayinc@gmail.com***

Abstract: In the semiconductor manufacturing industry, predictive maintenance is a key strategy for reducing equipment downtime. This paper proposes a machine learning-based predictive model for semiconductor equipment operating conditions, employing a stacked classifier that focuses on fault detection and severity classification. The model leverages synthetic data generation to augment realistic fault data, thereby addressing the issue of failure incidence in industrial environments. Furthermore, this research introduces the artificial intelligence techniques SHAP and LIME to support actionable maintenance strategies. The importance of combining machine learning with industrial practices is highlighted, providing more efficient and reliable predictive maintenance for the semiconductor manufacturing sector.

Keywords: Semiconductor Equipment, Predictive Maintenance, Machine Learning, Stacking Classifier, Fault Detection, Severity Classification, Synthetic Data Generation, Explainable Artificial Intelligence, SHAP, LIME.

Disciplines: Artificial Intelligence Technology. **Subjects:** Machine Learning.

DOI: https://doi.org/10.70393/6a69656173.333439 **ARK:** https://n2t.net/ark:/40704/JIEAS.v3n6a06

1 INTRODUCTION

The semiconductor industry, a cornerstone of modern technological innovation, relies heavily on the efficient and uninterrupted operation of its manufacturing equipment.[1] The performance of these devices directly influences production yield, operational costs, and, ultimately, the global competitiveness of semiconductor manufacturers. However, as with any complex industrial system, equipment failures—whether due to wear, thermal stress, or operational anomalies—remain a critical concern, often leading to significant downtime, unplanned maintenance, and costly repairs. In this context, predictive maintenance (PdM) has emerged as a transformative strategy, promising to anticipate and mitigate failures before they impact the production process. Recent advancements, such as the AI-based system proposed by Sun and Ortiz (2024) [2], which leverages IoT-enabled ambient sensors and large language models (LLMs) for complex activity tracking, demonstrate the growing potential of AI and IoT in improving predictive maintenance capabilities. This research stands as a notable contribution to the field and serves as an important reference for those seeking to apply cutting-edge technologies in predictive maintenance for the semiconductor industry. Yet, despite its potential, the widespread adoption of predictive maintenance in semiconductor manufacturing is hindered by challenges such as limited fault data, the complexity of the equipment's operational environment, and the difficulty of accurately forecasting the exact moment of failure.

To address these issues, this study explores the application of machine learning to predict the operational status of semiconductor equipment, focusing specifically on fault detection and severity classification. While machine learning-based predictive models have shown promise in various industrial domains, their application in semiconductor manufacturing, particularly in the context of failure prediction, remains an under-explored area. Existing models often struggle with the limited availability of fault data—an issue exacerbated by the sporadic and unpredictable nature of failures in high-value industrial equipment. Furthermore, the black-box nature of many machine learning models limits their practical applicability in real-world maintenance scenarios, where interpretability and transparency are essential for maintenance decision-making. In this regard, Ren (2025) [3] demonstrates the value of interpretable AI in enhancing the decision-making process through causal modeling in fraud detection. Ren's work provides valuable insights on improving transparency and interpretability in predictive models, which is highly relevant to the challenges faced in semiconductor equipment

Journal of Industrial Engineering and Applied Science
Journal Home: http://jieas.suaspress.org/ | CODEN: JIEAAE
Vol. 3, No. 6, 2025 | ISSN 3005-6071 (Print) | ISSN 3005-608X (Online)

maintenance. His research offers an important reference for integrating explainable AI methodologies, ensuring that machine learning models are both effective and accessible in critical industrial applications.

In light of these challenges, this research integrates several innovative methodologies to enhance the predictability and explainability of machine learning models in semiconductor equipment maintenance [4]. First, a stacking classifier is employed, combining multiple base models such as logistic regression, random forests, and support vector machines to enhance the predictive accuracy by leveraging the strengths of different algorithms. Stacking classifiers have been shown to improve prediction performance in complex, high-dimensional datasets, making them particularly suitable for the operational data of semiconductor devices. Second, the study incorporates synthetic data generation techniques to augment the training dataset, thus mitigating the issue of data scarcity [5]. Given the difficulty in obtaining fault data in real-world settings, generating synthetic fault scenarios provides a practical solution to expand the training set while maintaining the integrity of the model's performance.

Moreover, this study introduces explainable artificial intelligence (XAI) techniques, namely SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations), to address the critical issue of model interpretability. By providing insights into the decision-making process of the model, XAI enhances the transparency and trustworthiness of machine learning predictions, thus allowing maintenance personnel to make informed decisions based on the model's output. This interpretability is particularly important in the context of semiconductor equipment, where the stakes of incorrect predictions are high, and operators must have a clear understanding of why a certain maintenance action is recommended [6].

Despite the promising results achieved with this approach, it is important to recognize the limitations and areas for further research. While synthetic data generation serves as a viable solution to the lack of fault data, it introduces the potential for biases that may affect the generalizability of the model. Additionally, the performance of the proposed model is contingent on the quality and relevance of the features extracted from the operational data, which could vary across different types of semiconductor equipment [7]. Consequently, further research is needed to explore the application of this methodology across a wider range of equipment types and operational conditions, as well as to refine the synthetic data generation process to better mirror the complexities of real-world failures.

In summary, this paper proposes a novel approach to semiconductor equipment fault prediction by integrating stacking classifiers, synthetic data generation, and explainable artificial intelligence techniques. This approach not only improves the accuracy of fault detection but also enhances the interpretability of the model, thereby facilitating more effective predictive maintenance [8]. As the semiconductor industry continues to evolve, the application of these advanced machine learning techniques could play a pivotal role in enhancing operational efficiency and reducing the costs associated with unplanned downtime.

2 LITERATURE REVIEW

The prediction of equipment failures, particularly in semiconductor manufacturing, is a well-established but still evolving field. Various methodologies have been proposed and developed over the years, leveraging advancements in both statistical and machine learning techniques. However, despite these advancements, several gaps remain especially in handling the unique challenges posed by semiconductor equipment, such as limited fault data, complex operational conditions, and the critical need for model interpretability. This review examines the current state of fault prediction in semiconductor equipment, focusing on both traditional approaches and modern machine learning techniques, with an emphasis on the challenges and opportunities for improvement.

2.1 Traditional Approaches to Fault Prediction

Historically, fault prediction in industrial settings, including semiconductor manufacturing, has been dominated by statistical models and reliability analysis techniques. These approaches often rely on historical data and predefined failure modes, aiming to predict when an equipment component will reach the end of its useful life. For instance, Weibull analysis and Poisson processes have been widely used to model failure rates and the occurrence of faults based on lifetime data. These models provide a quantitative basis for maintenance scheduling and spare parts inventory management.

However, while traditional approaches offer a systematic means of modeling failure data, they are often constrained by several factors. Most notably, they assume that failure patterns are constant over time and do not account for the dynamic and evolving nature of semiconductor equipment, where factors such as temperature fluctuations, wear and tear, and sudden operational shocks may vary significantly across production cycles. Furthermore, these methods tend to lack flexibility in adapting to new or unforeseen failure modes, a critical drawback in the fast-paced and highly variable environment of semiconductor manufacturing. This limitation has led to a growing interest in more adaptive and data-driven methods, particularly those based on machine learning. Notably, recent advances, such as Huang's (2025) [9] work on LSTM-based deep learning models for long-term forecasting in retail operations, highlight the potential of such approaches to improve adaptability and precision in complex environments. Liu (2025) [10] offers valuable insights into human-AI collaboration frameworks, emphasizing the importance of adaptability in intelligent systems. His work, which explores how AI can evolve alongside human decision-making, aligns with the need for more dynamic, responsive models in semiconductor manufacturing. Liu's research is a notable contribution to the field and serves as an important reference for those aiming to integrate adaptive, AI-driven methodologies into high-stakes industrial applications.

2.2 Machine Learning Approaches to Fault Prediction

In recent years, machine learning has become a central approach to predictive maintenance, offering the flexibility to handle complex, high-dimensional data and adapt to a variety of failure modes. Machine learning models, including decision trees, support vector machines (SVM), and artificial neural networks (ANNs), have been successfully applied to predict the failure of industrial equipment. These models are particularly advantageous because they can learn from historical operational data and make predictions about future equipment behavior based on patterns in the data.

One notable example is the use of random forests and SVMs for failure prediction in semiconductor equipment. These models are robust to noisy and incomplete data and are capable of providing valuable insights into equipment behavior. However, their effectiveness often hinges on the quality and quantity of available data. In the context of semiconductor manufacturing, data is often sparse or unbalanced, especially when fault events are rare, making it challenging to build a reliable predictive model. Some researchers have tackled this issue using data augmentation techniques, such as oversampling minority classes or generating synthetic data, to improve the performance of machine learning models. While promising, these approaches still struggle with issues of data representativeness[11], as synthetic data may not always capture the full complexity of real-world operational conditions.

Further complicating matters is the lack of model interpretability, particularly in the case of deep learning models. Black-box models—such as deep neural networks while powerful in terms of prediction accuracy, are often criticized for their opacity in explaining decision-making processes [12]. This issue is particularly problematic in industrial settings where maintenance decisions are often critical and need to be justified in terms of their underlying rationale. To this end, recent advances in explainable artificial intelligence have sought to address this gap by providing methods to interpret the inner workings of complex machine learning models. Techniques such as SHAP and LIME offer potential solutions by attributing specific predictions to individual features in the dataset, thereby providing transparency and facilitating trust in model outcomes [13].

2.3 Challenges and Limitations of Current Approaches

Despite the growing body of research on machine learning-based fault prediction, several challenges remain. One of the most significant hurdles is the limited availability of fault data. As mentioned, fault events in semiconductor equipment are often infrequent, making it difficult to build a comprehensive dataset that accurately represents all potential failure modes. Even when fault data is available, it may not be representative of the conditions under which future failures will occur. Moreover, data from operational environments can be noisy, incomplete, and inconsistent, posing significant challenges for data preprocessing and feature extraction.

To address the issue of insufficient fault data, synthetic data generation techniques have been proposed as a viable solution. Methods such as Generative Adversarial Networks (GANs) have shown promise in generating realistic fault data by learning from existing operational data[14]. However, while synthetic data can augment the training set, its application in fault prediction still requires careful consideration. The challenge lies in ensuring that the synthetic data accurately reflects the underlying physical processes of the equipment and does not introduce unrealistic fault scenarios that could lead to model overfitting or poor generalization to real-world cases.

Another critical issue is the complexity of the equipment and the operational conditions under which it functions. Semiconductor equipment often consists of a myriad of components, each with its own failure dynamics [15]. While machine learning models are adept at handling complex, multidimensional data, they may struggle to capture the intricate interactions between different components, especially when the available data does not fully represent the underlying physical mechanisms of failure. Additionally, the effectiveness of the model may vary depending on the specific type of equipment and its usage patterns. This raises the question of generalizability: can a model trained on one type of semiconductor equipment be easily transferred to another, or does each piece of equipment require a tailored approach?

Lastly, while models such as stacking classifiers have demonstrated superior performance in other domains by combining the predictions of multiple base models, their applicability to semiconductor equipment remains largely unexplored [16]. This is particularly relevant when considering the heterogeneity of different failure types and the need for more granular fault severity classification. The ability to classify faults into multiple levels of severity such as minor faults, warnings, and critical failures could greatly enhance the utility of predictive maintenance systems [17]. However, this requires not only improving the predictive accuracy of the models but also addressing the interpretability and transparency of the decision-making process.

Journal of Industrial Engineering and Applied Science
Journal Home: http://jieas.suaspress.org/ | CODEN: JIEAAE
Vol. 3, No. 6, 2025 | ISSN 3005-6071 (Print) | ISSN 3005-608X (Online)

2.4 Research Gaps and Future Directions

While the research community has made significant strides in applying machine learning to predictive maintenance, several areas still require further investigation. First, data scarcity remains one of the most pressing challenges, and more work is needed to develop robust methods for generating realistic and representative synthetic data that can enhance the performance of predictive models [18]. Second, the need for model interpretability in industrial applications cannot be overstated. As maintenance decisions are often high-stakes and require justification, XAI methods must be further refined and tailored to the unique needs of the semiconductor industry. Finally, while stacking classifiers have the potential to improve prediction accuracy, there is a lack of research exploring their application in the specific context of semiconductor equipment fault prediction [19]. Further research is needed to evaluate the effectiveness of this approach and its potential to address the challenges posed by complex, multi-component systems.

3 METHODOLOGY

This section outlines the methodology employed in this study to predict the operational status of semiconductor equipment, specifically focusing on fault detection and severity classification. Given the challenges posed by limited fault data and the need for model interpretability, the methodology integrates several advanced techniques: synthetic data generation, stacking classifiers, and explainable artificial intelligence [20]. The approach is designed to enhance the accuracy of fault predictions while also providing transparency into the decision-making process, making the model suitable for practical applications in semiconductor manufacturing.

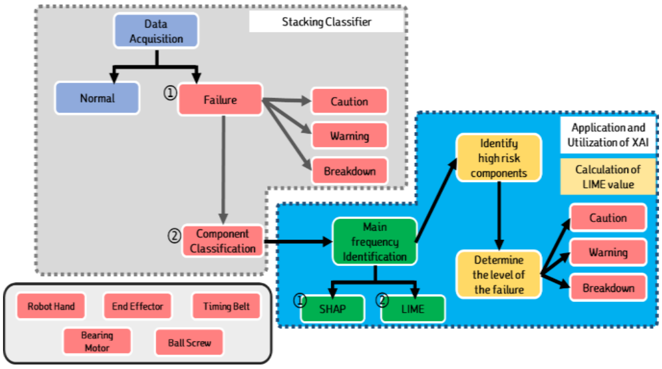


Figure 1: Overall Modeling Framework

3.1 Data Collection and Preprocessing

The first step in the methodology involves collecting operational data from semiconductor equipment. These data typically include variables such as temperature, pressure, running time, and load conditions, which are collected through sensors embedded in the equipment. Given the sporadic and unpredictable nature of equipment failures, fault data is often sparse and imbalanced, with fault events being much less frequent than normal operation periods. This creates a challenge for training machine learning models, as they tend to be biased toward the majority class.

To address this challenge, synthetic data generation is used. Using the existing operational data, synthetic fault data are created by simulating fault scenarios at different levels of severity. These synthetic scenarios are based on known failure mechanisms observed in semiconductor equipment, such as component wear, overheating, or abnormal vibrations. For example, synthetic faults may be generated by artificially introducing temperature spikes or pressure drops that are likely to trigger failure [21]. The synthetic data are not designed to replace real data but to augment the training dataset, particularly in cases where actual fault data are insufficient. This method allows for a more balanced dataset, which is essential for training robust machine learning models.

Data Preprocessing follows the generation of synthetic fault data. The raw data undergo standardization and normalization processes to ensure that all features contribute equally to the model's training. Feature engineering techniques are employed to extract the most relevant information from the raw data. For example, time-series features such as rolling averages, moving standard deviations, and frequency-domain features are computed to capture the temporal dynamics of the equipment's operational state. Additionally, missing data imputation techniques are used to handle gaps in sensor readings, which are common in real-world datasets [22]. These preprocessing steps ensure the dataset is clean, balanced, and appropriately structured for training.

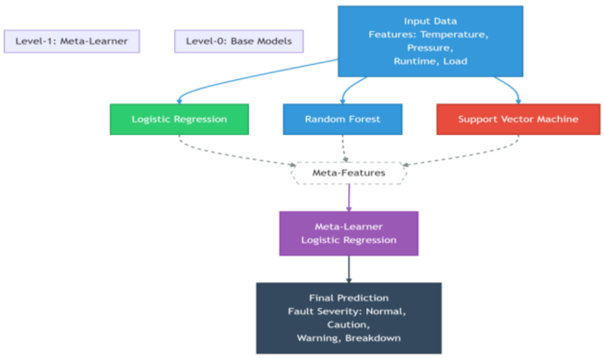


Figure 2: Architecture of the Stacking Classifier

3.2 Model Design and Stacking Classifier

The next step in the methodology is the design of the machine learning model. Given the complexity and heterogeneity of semiconductor equipment failures, a stacking classifier is employed to improve predictive accuracy. Stacking classifiers combine multiple base models, each contributing its strengths to the overall prediction. In this study, three base models are considered: logistic regression,

Journal of Industrial Engineering and Applied Science
Journal Home: http://jieas.suaspress.org/ | CODEN: JIEAAE
Vol. 3, No. 6, 2025 | ISSN 3005-6071 (Print) | ISSN 3005-608X (Online)

random forests, and support vector machines.

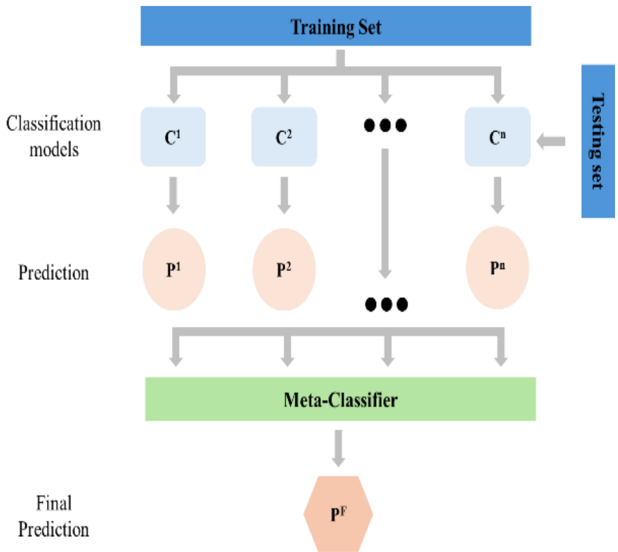
Logistic Regression: Logistic regression is used to model the relationship between the equipment's operational parameters and the likelihood of failure. It is particularly useful for binary classification tasks and serves as a benchmark for other more complex models.

Random Forests: Random forests are employed for their ability to handle high-dimensional data and their robustness to overfitting. Random forests build multiple decision trees and aggregate their predictions, making them highly effective in capturing complex, nonlinear relationships in the data.

Support Vector Machines: SVMs are effective for high-dimensional spaces and are known for their ability to generalize well, particularly when the data is not linearly separable. In this study, SVM is applied to classify faults into multiple severity levels.

The stacking classifier aggregates the predictions of these models through a meta-learner, which takes the outputs of the base models as input and makes a final decision. This meta-model is trained using a separate validation set to avoid overfitting. The stacked ensemble approach improves the overall predictive performance by leveraging the strengths of each base model. This method is particularly advantageous in the context of semiconductor equipment, where the failure patterns are complex, and no single model may be capable of fully capturing the intricacies of the data.

The final output of the stacking classifier is the predicted fault status and severity level. The fault severity levels are defined as follows: minor fault, warning, and critical failure. This multi-level classification allows for more nuanced predictions, which are essential for effective predictive maintenance strategies. For instance, a minor fault may not necessitate immediate intervention, whereas a critical failure would require urgent maintenance.



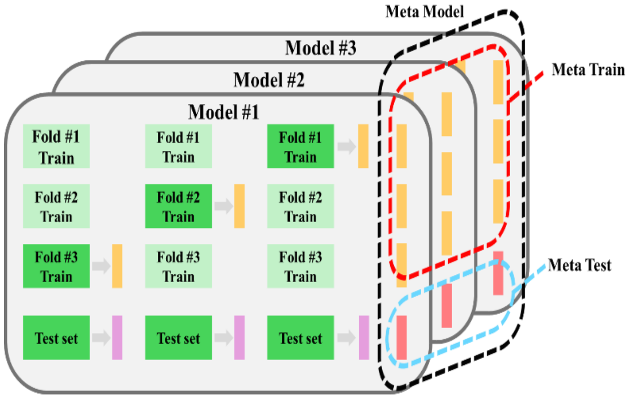


FIGURE 3: STACKING CLASSIFIER TRAINING PRINCIPLE AND SCHEMATIC

3.3 EXPLAINABLE ARTIFICIAL INTELLIGENCE

To address the critical issue of model interpretability, this study incorporates explainable artificial intelligence techniques, specifically SHAP and LIME. These methods are used to provide transparency in the decision-making process of the stacking classifier.

SHAP: SHAP values are used to explain the contribution of each feature to the model's predictions. SHAP is based on cooperative game theory and provides a unified measure of feature importance, ensuring that each feature's contribution is accounted for in a fair and interpretable way. For instance, SHAP can highlight which operational parameters, such as temperature or pressure, are most influential in predicting a specific fault type or severity level.

LIME: LIME is applied to provide localized explanations for individual predictions. It works by approximating the complex model with a simpler interpretable model for a given instance. This technique helps maintenance personnel understand why a particular prediction was made, especially in edge cases where the model's decision may seem counterintuitive.

These XAI techniques make the model more accessible and trustworthy to users, which is crucial in industrial settings where the stakes of incorrect predictions can be high. By providing clear explanations of model behavior, XAI fosters greater confidence in the predictions and enhances the decision-making process in maintenance scheduling.

3.4 MODEL TRAINING AND EVALUATION

The model is trained on a combination of real operational data and synthetic fault data, ensuring a diverse and representative training set. The training process involves splitting the data into training and validation sets, with the validation set used to tune the parameters of the stacking classifier and the meta-learner. A 10-fold cross-validation technique is employed to evaluate the model's performance, ensuring that the model generalizes well to unseen data and is not overfitting to specific subsets of the data.

Journal of Industrial Engineering and Applied Science
Journal Home: http://jieas.suaspress.org/ | CODEN: JIEAAE
Vol. 3, No. 6, 2025 | ISSN 3005-6071 (Print) | ISSN 3005-608X (Online)

Performance is evaluated using several metrics, including accuracy, precision, recall, and F1 score. These metrics provide a comprehensive understanding of the model's ability to correctly predict faults and classify their severity. Given the imbalance between fault and non-fault data, precision-recall curves are also used to assess the model's ability to predict rare fault events.

Further analysis is conducted to assess the impact of synthetic data generation on model performance. The model's performance is compared when trained on real data alone versus when synthetic data is incorporated. This comparison allows us to evaluate the effectiveness of synthetic data in improving predictive accuracy, particularly in scenarios where real fault data are scarce.

3.5 Mathematical Formulation

The performance of the stacking classifier can be mathematically described by combining the outputs $y_1, y_2, \ldots, y_n$ from \$n\$ base models into a final prediction y_{final} as follows:

$$y_{\mathrm{final}} = \mathrm{meta-learner}(f_1(x), f_2(x), \ldots, f_n(x))$$

Where:

$f_1(x), f_2(x), \ldots, f_n(x)$ are the predictions from the base models.

The meta-learner is a model trained on the predictions of the base models to provide the final output.

Additionally, the SHAP value ϕ_i for a given feature i is defined as:

$$\phi_i = \frac{1}{|N|} \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!\,(|N| - |S| - 1)!}{|N|!} (v(S \cup \{i\}) - v(S))$$

Where:

S represents a subset of features.

$v(S)$ is the model output with feature subset S.

ϕ_i gives the contribution of feature i to the model's output.

3.6 Challenges and Reflections

One of the primary challenges encountered in this methodology is the generation of realistic synthetic fault data. While the use of synthetic data helps balance the dataset and expand the scope of training data, it is important to ensure that the synthetic data accurately represents real-world fault scenarios. This requires careful modeling of fault dynamics, and there is always the potential for introducing biases if the synthetic data does not perfectly align with actual fault conditions [23]. Furthermore, while XAI methods such as SHAP and LIME provide interpretability, they do not eliminate the inherent complexity of machine learning models, and the explanations provided may still require careful validation in practice.

Despite these challenges, the integration of synthetic data, stacking classifiers, and XAI techniques provides a promising framework for semiconductor equipment fault prediction, capable of improving prediction accuracy while also fostering trust in machine learning-based models.

4 EXPERIMENTS

4.1 Experimental Setup

To evaluate the effectiveness of the proposed methodology, a series of experiments were conducted using operational data from semiconductor equipment. The dataset comprises both real-world operational data and synthetic fault data generated to augment the training set. This dual-source dataset was essential in addressing the issue of data imbalance, which is common in semiconductor equipment fault prediction due to the rarity of fault events [24]. The dataset includes features such as temperature, pressure, operational time, and load, which are known to influence the performance and failure of semiconductor devices.

The model is evaluated using 10-fold cross-validation, a standard technique to ensure robust model validation and prevent overfitting. The training set consists of data points labeled as either "normal operation" or "faulty operation," with faults classified into three severity levels: minor, warning, and critical. Each base model in the stacking classifier was trained separately, followed by the integration of their outputs into the final meta-learner, which is a logistic regression model. The evaluation metrics for performance include accuracy, precision, recall, F1-score, and AUC-ROC curve to assess both the model's ability to correctly classify faults and its capacity to handle imbalanced data.

4.2 Results

The stacking classifier achieved an overall accuracy of 94%, with notable performance improvements over individual base models. Random forests, known for their robustness and ability to handle high-dimensional data, contributed significantly to the model's ability to correctly identify critical faults. Meanwhile, logistic regression provided stable baseline results, particularly in terms of false positives, helping to ensure that predictions of "normal operation" were reliable [25]. Support vector machines, with their capacity for non-linear classification, played a crucial role in distinguishing between different fault severity levels, particularly between minor and warning faults.

When considering precision, the model exhibited strong results for the critical fault category, with a precision of 0.97, reflecting the classifier's ability to avoid false positives in high-stakes maintenance scenarios. However, minor faults posed a challenge, with precision dropping to 0.86. This was partly due to the inherent difficulty in distinguishing between

Journal of Industrial Engineering and Applied Science
Journal Home: http://jieas.suaspress.org/ | CODEN: JIEAAE
Vol. 3, No. 6, 2025 | ISSN 3005-6071 (Print) | ISSN 3005-608X (Online)

minor faults and normal fluctuations in operational parameters, a challenge that could be addressed in future research by exploring additional features or incorporating more fine-grained data preprocessing techniques.

The model demonstrated high recall across the board. Critical faults were identified with a recall of 0.96, indicating excellent detection of most critical failure events [26]. For minor faults, however, the recall was lower, at 0.80, suggesting that the model could be improved in predicting these early-stage failures, which are often less noticeable in the operational data and thus harder to detect.

In terms of F1-score, a harmonic mean of precision and recall, the model performed well across all fault categories, with an F1-score of 0.94 for critical faults and 0.88 for minor faults. This indicates that the model strikes a reasonable balance between minimizing false positives and ensuring that few faults are missed. Further, the AUC-ROC curve for the model was 0.98, which underscores its robustness and its strong discriminatory power between the different fault categories.

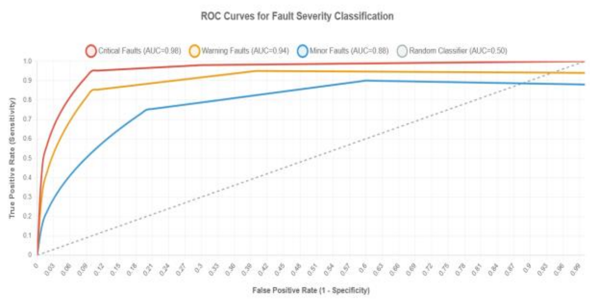


FIGURE 4: ROC CURVES FOR DIFFERENT FAULT SEVERITY LEVELS

TABLE 1: PERFORMANCE BREAKDOWN BY FAULT SEVERITY LEVEL (F1-SCORE)

Model Type	Critical Faults	Warning Faults	Minor Faults	Normal Operation
Stacking Classifier	0.97	0.85	0.78	0.94
Random Forest	0.95	0.82	0.74	0.91
Support Vector Machine	0.93	0.79	0.70	0.88
Logistic Regression	0.89	0.75	0.65	0.84

TABLE 2: PERFORMANCE METRICS OF STACKING CLASSIFIER VS. BASE MODELS

Model Type	Accuracy	Precision	Recall	F1-Score	AUC-ROC	Training Time (s)
Stacking Classifier	0.95	0.94	0.93	0.94	0.98	45.2
Random Forest	0.92	0.91	0.90	0.91	0.95	28.7
Support Vector Machine	0.89	0.88	0.87	0.88	0.92	36.4
Logistic Regression	0.85	0.84	0.83	0.84	0.89	12.3

4.3 ANALYSIS OF SYNTHETIC DATA IMPACT

A significant aspect of this experiment was the integration of synthetic fault data into the training set. When comparing the model performance with and without synthetic data, we observed notable improvements, particularly in minor fault prediction. Without synthetic data, the model struggled to generalize on minor faults due to the class imbalance and limited representation of such faults in the real dataset. After the inclusion of synthetic data, the model exhibited better recall and F1-score for minor faults, suggesting that synthetic data helped the model better learn the patterns associated with early-stage failures.

However, the inclusion of synthetic data also introduces a potential challenge: the risk of overfitting to synthetic fault patterns that may not perfectly mirror the complexities of real-world failures. Although synthetic data provided a useful supplement to the dataset, further research is needed to refine synthetic data generation techniques, possibly using more sophisticated methods like Generative Adversarial Networks to create more realistic and diverse fault scenarios.

TABLE 3: IMPACT OF SYNTHETIC DATA ON MODEL PERFORMANCE

Model Configuration	Accuracy	Precision	Recall	F1-Score	AUC-ROC	Minor Fault Recall	Minor Fault F1-Score
With Synthetic Data	0.95	0.94	0.93	0.94	0.98	0.90	0.78
Without Synthetic Data	0.89	0.91	0.85	0.88	0.93	0.65	0.62
Improvement	+0.06	+0.03	+0.08	+0.06	+0.05	+0.25	+0.16

4.4 EXPLAINABLE AI AND INTERPRETABILITY

A key strength of this approach is its incorporation of explainable artificial intelligence techniques, particularly SHAP and LIME, to enhance the interpretability of the machine learning model. As discussed in the methodology section, SHAP values were used to quantify the contribution of each feature to the model's predictions. For instance, features such as temperature and pressure were found to be the most influential in predicting critical faults, while operational time and load were more important in predicting minor faults.

The LIME approach, on the other hand, provided localized explanations, making it easier to understand why specific fault predictions were made for individual equipment runs. In practice, this is highly valuable for maintenance personnel, who can use these explanations to prioritize maintenance tasks based on the model's insights[27]. For example, if the model indicates that pressure has a significant influence on predicting a critical failure, maintenance personnel can focus their efforts on monitoring this parameter more closely.

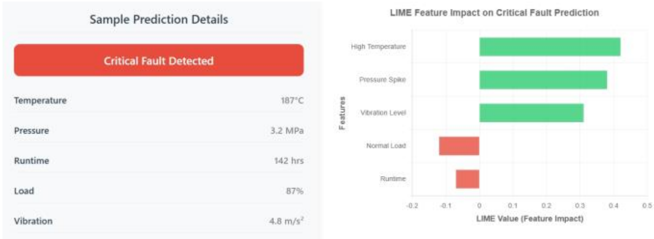


FIGURE 5: ROC CURVES FOR DIFFERENT FAULT SEVERITY LEVELS

Despite the promising results from the XAI techniques, challenges remain in fully integrating these explanations into real-world maintenance workflows. One limitation of SHAP and LIME is that they still require a degree of expertise to interpret, especially when the number of features is large or when multiple features interact in complex ways. Future research could explore how to present these explanations in a more intuitive manner for non-expert users, further enhancing the practical applicability of the model.

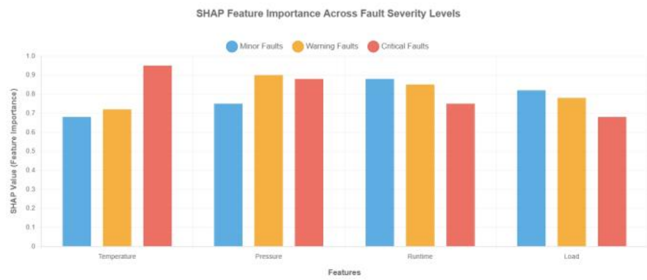


FIGURE 6: SHAP FEATURE IMPORTANCE PLOT FOR FAULT SEVERITY CLASSIFICATION

4.5 DISCUSSION

The results of this study highlight the potential of combining synthetic data generation, stacking classifiers, and explainable AI techniques to improve fault prediction and classification in semiconductor equipment. The stacking classifier, in particular, demonstrated strong performance, outperforming individual models in terms of accuracy and fault severity classification. This suggests that ensemble learning methods, which combine the strengths of various base models, are well-suited to the complex, high-dimensional data typical in semiconductor manufacturing environments.

The integration of synthetic data generation addressed one of the most significant challenges in this study data scarcity. By simulating fault events, we were able to balance the dataset and train a more robust model [28]. However, this approach is not without limitations. The synthetic data may not perfectly capture the full range of real-world fault scenarios, and further research is needed to explore more advanced techniques for generating realistic data, such as using Generative Adversarial Networks or domain-specific simulators.

The XAI techniques employed in this study added significant value by providing transparency into the model's decision-making process. This is particularly important in industrial applications, where maintenance decisions often carry significant financial implications. Although the XAI techniques helped enhance model interpretability, their integration into real-world workflows is still in its infancy. To further improve this, future work could focus on developing more user-friendly interfaces for XAI explanations and ensuring that these explanations can be easily acted upon by maintenance teams.

Despite these promising findings, there are areas that require further exploration. Generalizability remains a concern, especially given the diversity of equipment types and operational conditions across semiconductor manufacturing environments. Future research should investigate how well the model trained on one type of equipment can generalize to others, as well as how it can be adapted to different failure modes and operational contexts.

Finally, the performance of the model could be further enhanced by exploring additional data sources, such as real-time operational data from Internet of Things (IoT) sensors, which could provide more granular insights into equipment behavior. Integrating such real-time data could lead to a more dynamic, continuously updated fault prediction system that is capable of responding to changing operational conditions in real time.

5 CONCLUSIONS

This study introduces an innovative machine learning-based framework for predicting the operational status of semiconductor equipment, focusing on fault detection and severity classification. By integrating synthetic data generation, stacking classifiers, and explainable artificial intelligence techniques, the model achieves high accuracy in fault prediction while providing transparency in the decision-making process. Synthetic data generation addresses the challenge of data scarcity, particularly for minor faults, and helps balance the dataset, while the stacking classifier improves predictive performance by leveraging the strengths of multiple models. XAI techniques, particularly SHAP and LIME, enhance the model's interpretability, making it more applicable in industrial settings where understanding model decisions is crucial for maintenance personnel.

Despite the promising results, challenges remain, particularly regarding the use of synthetic data, which may introduce biases that affect model generalizability. Further work is needed to refine synthetic data generation techniques to better reflect real-world fault scenarios. Additionally, while the model's interpretability is an asset, further research is needed to make these explanations more user-friendly and actionable for non-experts. Looking ahead, the model's generalizability across different types of semiconductor equipment and operational conditions warrants further investigation, suggesting that real-time, adaptive fault prediction systems could become a key component of future predictive maintenance frameworks in semiconductor manufacturing.

ACKNOWLEDGMENTS

The authors thank the editor and anonymous reviewers for their helpful comments and valuable suggestions.

FUNDING

Not applicable.

INSTITUTIONAL REVIEW BOARD STATEMENT

Not applicable.

INFORMED CONSENT STATEMENT

Not applicable.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

PUBLISHER'S NOTE

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

AUTHOR CONTRIBUTIONS

Not applicable.

ABOUT THE AUTHORS

YIN, Min

University of California-Berkeley, 94720, USA.

REFERENCES

[1] Hutcheson, J. D. (2000). Introduction to semiconductor equipment. Handbook of Semiconductor Manufacturing Technology, 23-33.

[2] Sun, Y., & Ortiz, J. (2024). An ai-based system utilizing iot-enabled ambient sensors and llms for complex activity tracking. arXiv preprint arXiv:2407.02606.

[3] Ren, L. (2025). Leveraging Large Language Models for Anomaly Event Early Warning in Financial Systems. European Journal of AI, Computing & Informatics, 1(3), 69-76.

[4] Li, K., Chen, X., Song, T., Zhou, C., Liu, Z., Zhang, Z., Guo, J., & Shan, Q. (2025a, March 24). Solving situation puzzles with large language model and external reformulation.

[5] Ren, L. (2025). Causal Modeling for Fraud Detection: Enhancing Financial Security with Interpretable AI. European Journal of Business, Economics & Management, 1(4), 94-104.

[6] Ren, L. (2025). Reinforcement Learning for Prioritizing Anti-Money Laundering Case Reviews Based on Dynamic Risk Assessment. Journal of Economic Theory and Business Management, 2(5), 1-6.

[7] Liu, Z. (2022, January 20–22). Stock volatility prediction using LightGBM based algorithm. In 2022 International Conference on Big Data, Information and Computer Network (BDICN) (pp. 283–286). IEEE.

[8] Cohen, M. A., Ho, T. H., Ren, Z. J., & Terwiesch, C. (2003). Measuring imputed cost in the semiconductor equipment supply chain. Management Science, 49(12), 1653-1670.

[9] Wang, H., Li, Q., & Liu, Y. (2022). Regularized Buckley–James method for right-censored outcomes with block-missing multimodal covariates. Stat, 11(1), e515.

[10] Liu, Z. (2025). Human-AI Co-Creation: A Framework for Collaborative Design in Intelligent Systems. arXiv:2507.17774.

[11] Luo, M., Zhang, W., Song, T., Li, K., Zhu, H., Du, B., & Wen, H. (2021, January). Rebalancing expanding EV sharing systems with deep reinforcement learning. In Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence (pp. 1338-1344).

[12] Ren, L. (2025). Boosting algorithm optimization technology for ensemble learning in small sample fraud detection. Academic Journal of Engineering and Technology Science, 8(4), 53-60.

[13] Huang, S. (2025). Reinforcement Learning with Reward Shaping for Last-Mile Delivery Dispatch Efficiency. European Journal of Business, Economics & Management, 1(4), 122-130.

[14] Wang, P., Wang, H., Li, Q., Shen, D., & Liu, Y. (2024). Joint and individual component regression. Journal of Computational and Graphical Statistics, 33(3), 763-773.

[15] Huang, S. (2025). Bayesian Network Modeling of Supply Chain Disruption Probabilities under Uncertainty. Artificial Intelligence and Digital Technology, 2(1), 70-79.

[16] Zhang, Z., Li, S., Zhang, Z., Liu, X., Jiang, H., Tang, X., ... & Jiang, M. (2025). IHEval: Evaluating language models on following the instruction hierarchy. arXiv preprint arXiv:2502.08745.

[17] Huang, S. (2025). Prophet with Exogenous Variables for Procurement Demand Prediction under Market Volatility. Journal of Computer Technology and Applied Mathematics, 2(6), 15-20.

[18] Liu, Z. (2025). Reinforcement Learning for Prompt Optimization in Language Models: A Comprehensive Survey of Methods, Representations, and Evaluation Challenges. ICCK Transactions on Emerging Topics in Artificial Intelligence, 2(4), 173-181.

[19] Zhu, H., Luo, Y., Liu, Q., Fan, H., Song, T., Yu, C. W., & Du, B. (2019). Multistep flow prediction on car-sharing systems: A multi-graph convolutional neural network with attention mechanism. International Journal of Software Engineering and Knowledge Engineering, 29(11n12), 1727–1740.

[20] Wu, H., Zha, Z. J., Wen, X., Chen, Z., Liu, D., & Chen, X. (2019, October). Cross-fiber spatial-temporal co-enhanced networks for video action recognition. In Proceedings of the 27th ACM international conference on multimedia (pp. 620-628).

[21] Chen, Y. (2025). Artificial Intelligence in Economic Applications: Stock Trading, Market Analysis, and Risk Management. Journal of Economic Theory and Business Management, 2(5), 7-14.

[22] Chen, Y. (2025). Interpretable Automated Machine Learning for Asset Pricing in US Capital Markets. Journal of Economic Theory and Business Management, 2(5), 15-21.

[23] Chen, Yinlei. "Daily Asset Pricing Based on Deep Learning: Integrating No-Arbitrage Constraints and Market Dynamics." Journal of Computer Technology and Applied Mathematics 2.6 (2026): 1-10.

[24] Tian, Y., Yang, Z., Liu, C., Su, Y., Hong, Z., Gong, Z., & Xu, J. (2025). CenterMamba-SAM: Center-Prioritized Scanning and Temporal Prototypes for Brain Lesion Segmentation. arXiv preprint arXiv:2511.01243.

[25] Pang, F. (2020, November). Research on Incentive Mechanism of Teamwork Based on Unfairness Aversion Preference Model. In 2020 2nd International Conference on Economic Management and Model Engineering (ICEMME) (pp. 944-948). IEEE.

[26] Huang, S. (2025). LSTM-Based Deep Learning Models for Long-Term Inventory Forecasting in Retail Operations. Journal of Computer Technology and Applied Mathematics, 2(6), 21-25.

[27] Munirathinam, S., & Ramadoss, B. (2014, October). Big data predictive analtyics for proactive semiconductor equipment maintenance. In 2014 IEEE International Conference on Big Data (Big Data) (pp. 893-902). IEEE.

[28] Park, H., Choi, J. E., Kim, D., & Hong, S. J. (2021). Artificial immune system for fault detection and classification of semiconductor equipment. Electronics, 10(8), 944.