Cloud-Based Watchdog Framework for Remote Automatic Correction of AI Models in CNC Machine Fault Prediction

SAI PRIYANKA KATABATHUNI
skatabat92@gmail.com

Abstract: Smart systems have assisted manufacturers to predict and prevent machine breakdown, which reduces downtime and maintenance expenses. Ensuring that AI models are continuously corrected without human manual intervention by means of a cloud-based infrastructure and a continuously running WATCHDOG module. We check fresh data sets, update the predictive model if performance drops and automatically deploy the best version to all user platforms. The system is trained to predict the CNC machine tool wear by the XGBoost algorithm and classify the results of the inspection as pass or fail. To enhance the accuracy of the model, a hybrid RF–XGBoost algorithm with CHI-Squared feature selection technique is used for retraining.A retraining process using the hybrid approach of the RF–XGBoost algorithm and a feature selection technique CHI-Squared is used to improve the accuracy of the model. Experimental results for the CNC tool wear data set demonstrate 100% prediction accuracy, which is better than the baselines. The platform leverages automated, continuous integration & deployment, web-based model evaluation and secure cloud data management for security and efficiency. This approach is more flexible, reduces computational costs for maintenance and offers real time, high accuracy industrial defect detection.

*“**Index Terms:** Automated deployment, automated inte gration, cloud service, dynamic model correction, model monitoring, smart manufacturing”.*

1. INTRODUCTION

AI is a key technology driving innovation across numerous industries in the quickly changing digital landscape. AI is being used more in manufacturing to improve production efficiency, predictive maintenance, and process optimization [1]. As the era of smart manufacturing gets here, information-driven decision-making and adaptive operational control are reshaping industrial ecosystems from the normal automation era. With the help of cloud computing and IIoT, enterprises can access huge amounts of sensor data to train intelligent systems to detect faults, optimize operations and save downtime [3]. These innovations have enhanced the quality of products, sustainability of production as well as saving energy [4]. However, many companies find it difficult to incorporate AI into their existing manufacturing systems.

Complex model management, time to deployment and lack of flexibility to changing production conditions hinder industrial AI installations [5]. But with traditional cloud solutions, developers are required to participate, take a lot of time to configure and to maintain it onsite, which raises costs and restricts scalability [6]. Private and sensitive information is sometimes included in industrial data when it is transferred to the cloud servers, creating privacy and network security issues [7]. However, many small to medium sized companies are not able to manage and update predictive models, limiting their ability to fully leverage AI technologies. Therefore, an automated, safe, and unified framework of deployment, monitoring, and real-time adaptation of AI models in industrial environments is required [8].

This work introduces an intelligent cloud-based system architecture that can automatically correct, retrain and deploy AI models in distributed industrial environments. The proposed system aims to dynamically modify the models based on operational data in a continuous learning framework to minimize the involvement of human in the system. It also offers real-time access to AI services through the web and an industrial secure communication architecture to safeguard industrial data [9]. The comprehensive and scalable AI integration platform can help enterprises minimize downtime, accelerate model updates, and enhance production reliability.

This can be considered as a paradigm shift for implementing intelligent models in industrial applications [10]. The platform offers enterprises a transition from automation to smart manufacturing, with its features of autonomous model optimization and secure remote access. The result is improved manufacturing efficiencies, decision-making accuracy and global industry sustainability and competitiveness.

2. LITERATURE REVIEW

AI, cloud computing, and automation have influenced the industrial digitalization. There are a number of research works that focus on enhancing the performance, scalability and dependability of cloud-based AI systems. Westin [11] showed that containerization can be used to decrease the computational overhead and resource waste in virtualized systems. While container technologies could be used to improve energy efficiency, the objectives of this study were geared toward optimizing the infrastructure level, instead of smart automation of AI model management. Jenkins pipeline visualization tools have been developed by Révész and Pataki [12] to make continuous integration workflow more transparent. Their approach enabled the monitoring and debugging of the system, but failed to handle industrial dynamic AI model correction.

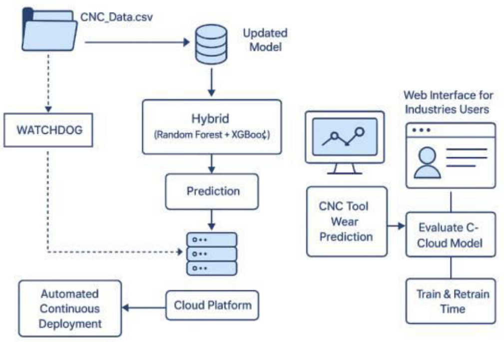
There has been research into enhancing collaborative management and automatic integration. Biesuz et al. [13] introduced HOG, an extension to version-control and distributed development in hardware design repositories by using Git-based collaboration. It succeeded in code management, but not for AI model retraining and automated deployment. Kim et al. [14] minimised the errors in data collection in the ecosystem and the manual effort on continuous integration through GitHub Actions. This approach only worked for data processing pipelines, neither autonomous AI model correction or industrial real-time deployment.

Shih et al. [15] demonstrated the feasibility of the cloud-based AI solution for air quality monitoring and prediction.The Azure based air quality early warning system (AQEWS) by Shih et al. [15] proved that cloud-based AI solutions for air quality monitoring and prediction are possible. The technology was scalable and available but did not have any self-correcting feature of changing data sets. Lai et al. [16] has been proposed an ensemble ML framework based on commercial gas sensors for the air pollution identification. They demonstrated good detection accuracy and sensor adaptation but neither continuous learning or remote retraining in distributed networks. According to Tung [17] virtual microservice-based architectures offer flexible update of AI models as they are deployed remotely. It was based on system architecture and did not validate the experimental system in real time industrial applications.

ML theory and practice have influenced Industrial AI Systems. Jordan and Mitchell [18] provided a solid overview of the trends and prospects of ML, noting that the need for scalable and generalizable models that could be adapted to diverse data sources was paramount. Sun [19] proposed the principles of these and developed the CNC Mill Tool Wear dataset to predict the condition of manufacturing tools. The dataset is vital for assessing the performance and robustness of failure prediction system models, such as XGBoost, RF and DT, which are excellent for classification, as pointed by Çetinkaya [20] without considering automated retraining or cloud deployment.

3. MATERIALS AND METHODS

The proposed intelligent approach in cloud computing automatically corrects, retrains and deploys AI models in industrial machine malfunction prediction. The data is from the Kaggle dataset, Tool Wear Detection in CNC Machines, and consists of sensor readings with the labels being outputted from the tool inspection. WATCHDOG module is continuously running to track performance degradation and automatically adjust models during data import, data preprocessing, model training and dynamic retraining. The predictive engine uses XGBoost as the main learner, augmented by CHI-Squared feature selection to extract the most relevant features and reduce computational redundancy. Robustness and generalization of prediction is increased by using a Hybrid ensemble model based on RF and XGBoost. Secure data upload, model evaluation, and real-time failure prediction are possible with the framework's cloud infrastructure and Flask-based web interface. This technique guarantees scalability, adaptability, accuracy, and continual optimization in distributed industrial situations.



“Fig.1 Proposed Architecture”

The proposed intelligent framework in the cloud is an integrated system that includes a continuously running WATCHDOG module for collecting, preprocessing, training, testing and auto-re-training the data. A secure system for dynamic AI model correction and deployment across distributed industrial users in the cloud. Feature selection, hybrid ensemble learning and automated CI/CD pipelines allow for continuous optimization. Users can upload machine data, monitor performance, and get real-time failure prediction with an encrypted data storage and sync capability via a Flask web interface.

a) Dataset Collection:

The CNC Tool Wear Dataset is obtained from UCI ML Repository and is used as an experimental dataset for this study. This dataset comprises of 10,000 samples of CNC parts produced under different process conditions. There are 21 numerical and 21 category features per sample, such as spindle speed, feed rate, cutting force, vibration signals and temperature. The class labels of the data set are binary: either “Pass” (no wear) or “Fail” (wear was detected). It is very imbalanced and has realistic industry data distribution, which makes the data difficult for the model to be generalized. It is representative of actual industrial applications due to the variety of materials and machining conditions used. This rich data set and structured labelling is perfect for testing predictive maintenance algorithms and comparing adaptive AI model correction systems in cloud-based industrial platforms.

b) Modules:

Watchdog: The Watchdog module evaluates incoming data sets for accuracy of the AI models. After copying the file, CNC_Data.csv, to WatchData.csv, users execute runWatchDog.bat. The module is updated with new data for training the XGBoost model and the data is deleted. The module automatically retrains the model when there is a decrease in the amount of data that is fed into the model, to increase accuracy. For industrial use, the new corrected model can be easily downloaded, as all modified model weights are stored in the cloud. Continuous adaptive learning without real-time streaming data is achieved via this simulation.

Web Interface for Industry Users: Industrial users can interact with the cloud-based AI system from a central point via the online interface. It enables the customers to submit machine sensor data, handle accounts and view predictions. The user can check the model performance, monitor training process, and receive the real-time update information via the interface. The UI streamlines the operations and ensures the security of the cloud data and model weights. Easy to use and efficient, it offers numerous functions for industrial users to communicate with the automated AI prediction system.

New Platform User Signup: The sign-up module is used for registering new industrial users in the cloud platform. Users enter their name, email, logon information. The system checks the input and establishes secure accounts to verify it. Users that have registered can access model evaluation, prediction and monitoring. This module regulates access to the system, to prevent unauthorized users. The initial step to cloud services makes it effortless to become involved with automatic model adjustment, real-time prediction and performance monitoring.

User Login: A safe login module enables registered users to log in to their accounts. Secure authentication is necessary to enable the use of platform features. After login, users can log machine sensor data, assess the performance of the model, and observe the prediction results. For access control, the system performs identification and session integrity checks to prevent un-authorised access. This module enables industrial customers to safely and with limited access to the cloud, use automatic model correction efficiently. Unlocks all platform interactive and analytical tools.

Evaluate Cloud Model: This module provides users a way to assess a cloud-based AI model. Users can measure the accuracy, precision, recall and F-Score to gauge the performance of the system in predicting. The module collects the weights of models in the cloud, evaluates predictions on the datasets provided by the user and displays the results online. It provides a live analysis on the effectiveness of the model, to decide retraining or updates. The module brings automated industrial predictions under the spotlight and makes them reliable.

Train & Retrain Time: The Train & Retrain Time module will display the time it takes for the AI model to be trained and retrained. Users can check the time spent in processing the datasets and updating the model's weights. The module compares efficiency and performance of incremental updating with of initial training. Visualizing the training time shows the needs of computation and responsiveness of the systems. The functionality allows automatic model correction to be transparent, enabling industrial users to understand the operational timeframes and monitor the learning of the AI model on the cloud.

CNC Tool Wear Prediction: The CNC Tool Wear Prediction module allows industrial customers to share machine sensor test data with the prediction result. The AI model analyzes the data input and makes calculations regarding tool inspection and maintenance. The online interface shows predictions to aid with decision-making for machine operation. For cloud-based models, the module automatically applies weights from the latest retraining rounds, if available. It gives real-time machine condition insights, decreasing downtime and maintenance costs and improving operational efficiency and reliability.

c) Algorithms:

XGBoost: XGBoost is an improved version of the Gradient Boosting framework that is used in supervised learning to improve the accuracy of prediction in addition to computational efficiency. It sequentially creates an ensemble of decision trees, rectifying residual error of each tree. Regularization helps avoid overfitting, parallel computing reduces the time to calculate, and pruning the trees enhances structure and performance. It is a scalable, generalizable and long-lasting system, where fault prediction and adaptive learning is possible in dynamic situations.

$$\hat{y}_i = \sigma\left(\sum_{k=1}^{K} f_k(x_i)\right), f_k \in F \quad (1)$$

Hybrid with Optimized Features (RF + XGBoost): RF + XGBoost ensemble is a hybrid model with greater predictive stability and accuracy. RF does this by stacking bootstrapped trees and XGBoost does this by boosting the trees sequentially. The hybrid model minimizes data redundancy and noise due to using statistically relevant feature subsets. This fusion boosts model robustness, flexibility, and reliability, ensuring high performance predictions in complex and unpredictable operation environments.

4. EXPERIMENTAL RESULTS

Accuracy: The accuracy of a test is its ability to distinguish patient from healthy cases. Measure test accuracy by computing the proportion of true positive and true negative for all cases that are tested. Mathematically, this is:

$$Accuracy = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \, (2)$$

Precision: Precision is the percentage of positive cases or sample correctly classified. The precision is computed as:

$$Precision = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \, (3)$$

Recall: The ability of a model to find all relevant examples of a class is assessed through the ML recall. It compares the number of accurate predictions of positive instances to the set of all instances of the class to demonstrate the model's completeness in capturing instances.

$$Recall = \frac{\text{TP}}{\text{TP} + \text{FN}} \, (4)$$

F1-Score: The accuracy of the ML models is calculated by F1 score. Caching both precision and recall scores for the models. The accuracy statistic represents the number of times the model was accurate over the entire data set.

$$F1\ Score = 2 * \frac{Recall \text{ X } Precision}{Recall + Precision} * 100 (5)$$

Table.1 Performance Evaluation Table

Algorithm Name	Accuracy	Precision	Recall	Fscore
Propose XGBoost	99.699	99.758	99.603	99.68
Extension Hybrid with Optimized Features	**100.0**	**100.0**	**100.0**	**100.0**

Table.1 highlights that the Hybrid with Optimized Features model outperforms the recommended XGBoost model in terms of Accuracy, Precision, Recall, and F-score.

Fig.2 Comparison Graph

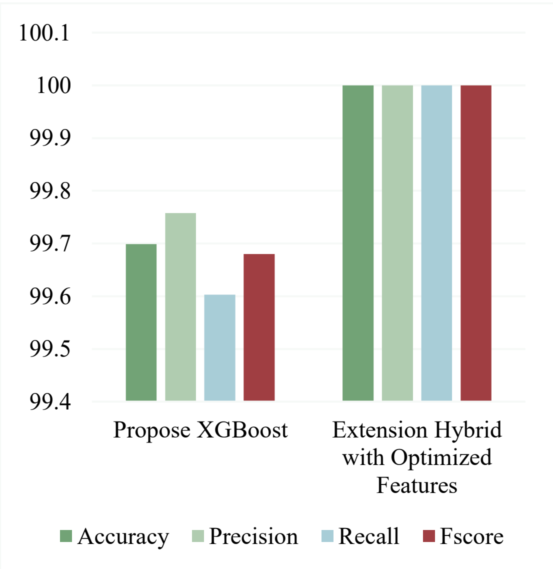
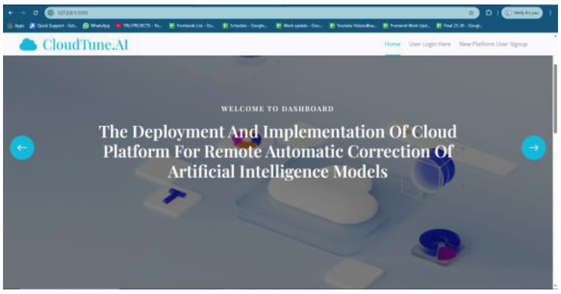
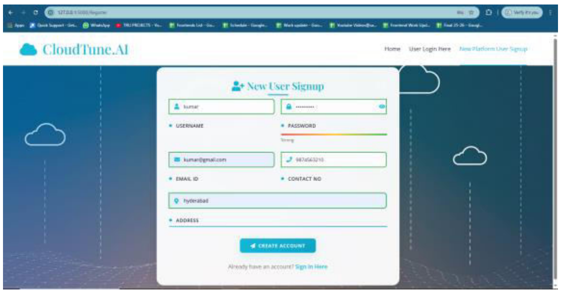


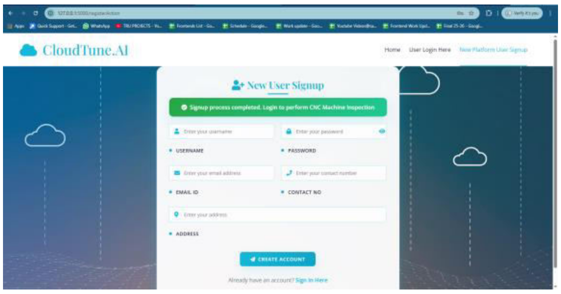
Fig.2, the Hybrid with Optimized Features model achieves higher scores on all the evaluation metrics like Accuracy, Precision, Recall and F-Score when compared to the suggested XGBoost.



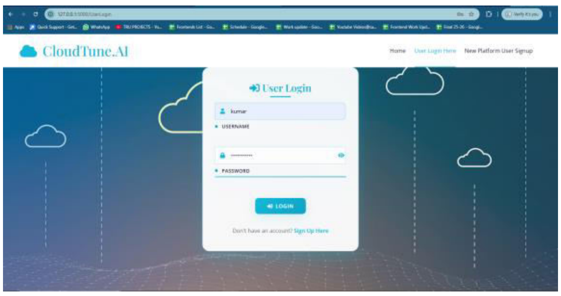
Click the ‘New User Sign up’ above to go to below page.



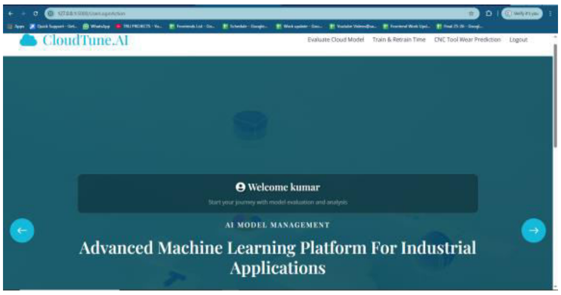
Above screen is showing users entering sign up information and pressing a button to navigate to the next page.



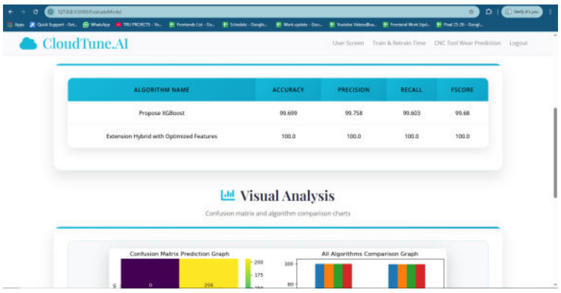
On above screen user sign up is completed, click on ‘User Login' to see below page.



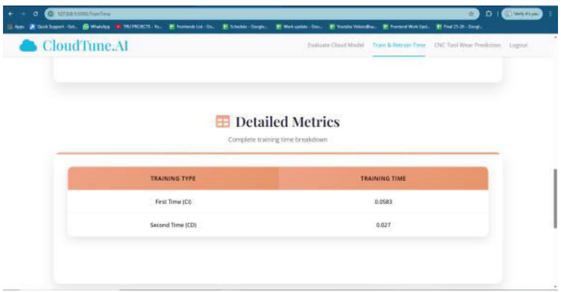
The above screen is a login screen and the below screen is the screen after logging in.



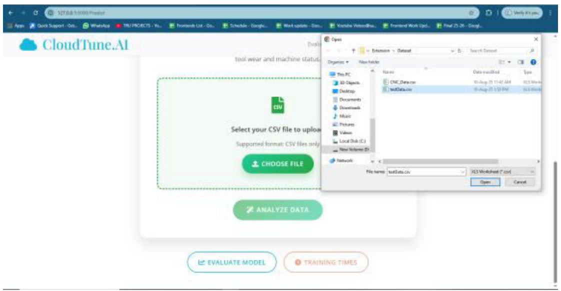
Click 'Evaluate Model' above to see algorithm performance then below.



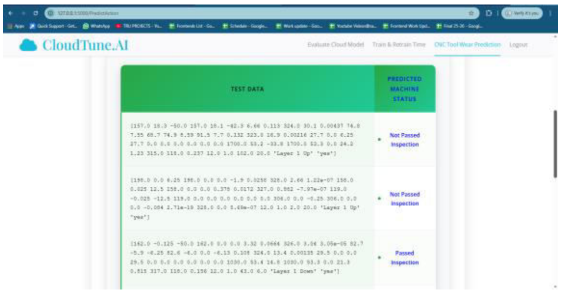
The tabular screen above displays the accuracy of XGBOOST to be 99% for the proposal. Following features selection and extension with Hybrid XGBOOST, accuracy was 100%. Precision, recall and FSCORE are presented in the table above. The x-axis in confusion matrix graph denotes 'Predicted Labels' and y-axis denotes True Labels. Green and yellow boxes indicate number of correct predictions, and blue boxes indicate a few incorrect predictions. Second graph shows algorithm performance based on algorithm names on the x-axis and accuracy and other metrics in colour bar on the y-axis. Extension Hybrid was correct for all of the algorithms. Click 'Train & Retrain Time' for watchdog training time.



Above screen displays watchdog training time (first/second run). Click on 'CNC Tool Wear Prediction' to view the following page.



From above screen user can click buttons to move to the below screen, select and upload test data file.



The first column indicates Machine Test results and the second column indicates whether the machine will pass Inspection in the above screen.

5. CONCLUSION

The cloud-based remote correction framework makes AI model refinement and deployment in industrial fault prediction easier and more convenient. WATCHDOG module features a smooth update without developers' intervention and enables dynamic model monitoring as well as automatic retraining. All cloud platform users have real-time access to optimized model performance with this technique. The platform achieved 100% peak accuracy with both the Hybrid RF – XGBoost algorithm and CHI-Squared feature selection, which was a significant improvement over the accuracy of the basic XGBoost model of 99.39%. The improvements demonstrate the reliability of the predictions for machine tool wear detection, based on automated model correction and optimization. It also features secure data management, automated continuous integration and deployment, and an easy-to-use online interface for evaluating data in real-time. Averaged out, the technique cuts down on upkeep expenses and reliance on the developer, and guarantees self-correcting adaptive savvy. This innovation is a major step toward completely autonomous industrial AI systems with great precision and stability in different operating settings.

The integration of the multi-sensor data with the cloud infrastructure will be expanded to handle large-scale diversified environments at the industrial scale. The DL architectures, such as CNNs and RNNs, will be added to deal with complex temporal and spatial patterns in machine data. We'll explore techniques for adaptive hyperparameter optimization and automated feature engineering to enhance the accuracy and generalization of the model. Real-time data streaming and predictive maintenance across distributed systems will be possible with IoT frameworks. A self-adaptive security layer and an audit system based on blockchain would ensure the transparency, reliability, and integrity of data during updates and when decisions are made in the cloud.

REFERENCES

[1] Kiangala, K. S., & Wang, Z. (2020, August). An adaptive framework for configuration of parameters in an Industry 4.0 manufacturing SCADA system by merging machine learning techniques. In 2020 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD) (pp. 1-6). IEEE.

[2] Sinha, S., & Lee, Y. M. (2024). Challenges with developing and deploying AI models and applications in industrial systems. Discover Artificial Intelligence, 4(1), 55.

[3] Javaid, M., Haleem, A., Singh, R. P., & Suman, R. (2022). Artificial intelligence applications for industry 4.0: A literature-based study. Journal of Industrial Integration and Management, 7(01), 83-111.

[4] Wang, X., Hu, H., Wang, Y., & Wang, Z. (2024). IoT real-time production monitoring and automated process transformation in smart manufacturing. Journal of Organizational and End User Computing (JOEUC), 36(1), 1-25.

[5] Chang, R. I., Lee, C. Y., & Hung, Y. H. (2021). Cloud-based analytics module for predictive maintenance of the textile manufacturing process. Applied Sciences, 11(21), 9945.

[6] S. Kinkel, M. Baumgartner, and E. Cherubini, "Prerequisites for the adoption of AI technologies in manufacturing– Evidence from a world wide sample of manufacturing companies," Technovation, vol. 110, 2022, Art. no. 102375.

[7] H. Ding, R. X. Gao, A. J. Isaksson, R. G. Landers, T. Parisini, and Y. Yuan, "State of AI-based monitoring in smart manufacturing and introduction to focused section," IEEE/ASME Trans. Mechatron., vol. 25, no. 5, pp. 2143–2154, Oct. 2020, doi: 10.1109/TMECH.2020.3022983.

[8] P. R. Desai, "A survey of performance comparison between virtual machines and containers," Int. J. Comput. Sci. Eng., vol. 4, no. 7, pp. 55–59, 2016.

[9]R.R.Yadav,E.T.G.Sousa,andG.R.A.Callou,"Perf ormancecomparison between virtual machines and Docker containers," IEEE Latin Amer. Trans., vol.

16, no. 8, pp. 2282–2288, Aug. 2018, doi: 10.1109/TLA. 2018.8528247.

[10] H.Goldstein, "Police discretion: The ideal versus the real," Public Admin. Rev., vol. 23, no. 3, pp. 140–148, 1963.

[11] J. Westin, "Evaluation of energy consumption in virtualization envi ronments: Proof of concept using containers," Bachelor's thesis, Dept. Comput. Sci., Umeå Univ., Umea, Sweden, 2017.

[12] Á. Révész and N. Pataki, "Visualisation of Jenkins pipelines," Acta Cy bernetica, vol. 25, no. 4, pp. 877–895, 2022.

[13] N. Biesuz, A. Camplani, D. Cieri, N. Giangiacomi, F. Gonnella, and A. Peck, "Hog (HDL on git): A collaborative management tool to handle git-based HDL repository," J. Instrum., vol. 16, no. 4, 2021, Art. no. T04006.

[14] A. Y. Kim et al., "Implementing GitHub actions continuous integration to reduce error rates in ecological data collection," Methods Ecol. Evol., vol. 13, no. 11, pp. 2572–2585, 2022.

[15] D.-H. Shih, T.-W. Wu, W.-X. Liu, and P.-Y. Shih, "An azure ACES early warning system for air quality index deteriorating," Int. J. Environ. Res. Public Health, vol. 16, no. 23, 2019, Art. no. 4679.

[16] W.-I. Lai, Y.-Y. Chen, and J.-H. Sun, "Ensemble machine learning model for accurate air pollution detection using commercial gas sensors," Sen sors, vol. 22, no. 12, 2022, Art. no. 4393.

[17] T.-Y. Tung, "Virtual microservice technology based innovative platform architecture development for AI model dynamic modification and remote automated deployment," Master's thesis, Inst. Manuf. Inf. Syst., National ChengKungUniv.,Tainan, Taiwan, 2022. [Online]. Available: https://hdl. handle.net/11296/u6etpp

[18] M.I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," Science, vol. 349, no. 6245, pp. 255–260, 2015.

[19] S. Sun, "CNC mill tool wear," kaggle.com, 2018. [Online]. Available: https://www.kaggle.com/datasets/shasun/tool-wear-detection-in-cnc mill

[20] E. Ç. Etinkaya, "(XGBoost & RanFor & DecTrees) basic training," kag gle.com, Nov. 2021. [Online]. Available: https://www.kaggle.com/code/ esragcetinkaya/xgboost-ranfor-dectrees-basic-training

[21] D. Minovski, C. ˚ Ahlund, K. Mitra, and R. Zhohov, "Defining qual ity of experience for the Internet of Things," IT Prof., vol. 22, no. 5, pp. 62–70, 2020.