



## Digital Twin-Based Automated Fault Diagnosis in Industrial IoT Applications

Samah Alshathri<sup>1</sup>, Ezz El-Din Hemdan<sup>2</sup>, Walid El-Shafai<sup>3,4,\*</sup> and Amged Sayed<sup>5,6</sup>

<sup>1</sup>Department of Information Technology, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh, 11671, Saudi Arabia

<sup>2</sup>Department of Computer Science and Engineering, Faculty of Electronic Engineering, Menoufia University, Menouf, 32952, Egypt

<sup>3</sup>Computer Science Department, Security Engineering Lab, Prince Sultan University, Riyadh, 11586, Saudi Arabia

<sup>4</sup>Department of Electronics and Electrical Communications Engineering, Faculty of Electronic Engineering, Menoufia University, Menouf, 32952, Egypt

<sup>5</sup>Department of Industrial Electronics and Control Engineering, Faculty of Electronic Engineering, Menoufia University, Menouf, 32952, Egypt

<sup>6</sup>Department of Electrical Energy Engineering, College of Engineering & Technology, Arab Academy for Science Technology & Maritime Transport, Smart Village Campus, Egypt

\*Corresponding Author: Walid El-Shafai. Email: eng.waled.elshafai@gmail.com

Received: 05 July 2022; Accepted: 31 October 2022

**Abstract:** In recent years, Digital Twin (DT) has gained significant interest from academia and industry due to the advanced in information technology, communication systems, Artificial Intelligence (AI), Cloud Computing (CC), and Industrial Internet of Things (IIoT). The main concept of the DT is to provide a comprehensive tangible, and operational explanation of any element, asset, or system. However, it is an extremely dynamic taxonomy developing in complexity during the life cycle that produces a massive amount of engendered data and information. Likewise, with the development of AI, digital twins can be redefined and could be a crucial approach to aid the Internet of Things (IoT)-based DT applications for transferring the data and value onto the Internet with better decision-making. Therefore, this paper introduces an efficient DT-based fault diagnosis model based on machine learning (ML) tools. In this framework, the DT model of the machine is constructed by creating the simulation model. In the proposed framework, the Genetic algorithm (GA) is used for the optimization task to improve the classification accuracy. Furthermore, we evaluate the proposed fault diagnosis framework using performance metrics such as precision, accuracy, F-measure, and recall. The proposed framework is comprehensively examined using the triplex pump fault diagnosis. The experimental results demonstrated that the hybrid GA-ML method gives outstanding results compared to ML methods like Logistic Regression (LR), Naïve Bayes (NB), and Support Vector Machine (SVM). The suggested framework achieves the highest accuracy of 95% for the employed hybrid GA-SVM. The proposed framework will effectively help industrial operators make an appropriate decision concerning the fault analysis for IIoT applications in the context of Industry 4.0.



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Keywords:** Automated fault diagnosis; control system; ML; AI; CC; IIoT; digital twins; genetic algorithm; GA-ML technique

## 1 Introduction

Recently, the evolution in information technologies such as the Internet of Things (IoT), Cloud Computing (CC), and Cyber-Physical Systems (CPSs), besides development in communication technologies, are revolutionizing the system approach to transmitting information between various sources. Undoubtedly, the revolution in all aspects of life nowadays is due to digitalization. Because of this advance and rebellion in Information Technology and Industry 4.0, the Digital Twins (DT) has come to be one of the talented ideas [1–8]. The DT is a digital model of real products from the perspective of CPSs. Therefore, a virtual complement of a convinced system mimics its real behavior. During full lifecycle development, the digital data from the virtual system, which is coupled to a real system, characterizes the information from the physical system. Consequently, integrating digital with its physical twins makes it an efficient method to handle, control and support the decision-making when the real system is running [9]. The DT records all data collected from physical sensors as well as the execution system response. As a result, the critical role of DT is to predict and diagnose the behavior of the physical system to foresee malfunction or fault, and feed data to the system to provide the best possible maintenance [10].

The virtual representation of the system is the major goal of DT to enhance productivity, increase operation flexibility, and decrease maintenance time and costs. On the other hand, DT can be run depending on the requirement, either on a cloud-hosted system or at the edge layer, to address the issues of Industry 4.0, which arise from the production of linked components. DT, on the other hand, supports both a clarification of the behavior of the actual system and the best possible solutions for the physical model. To improve the control action, foresee the system performance, and endorse decision making, DT uses a fundamental modeling system, simulation procedures, and transparent simulation [11–14].

The digital twin concept was first introduced by Grieves in 2002 [15], and then National Aeronautics and Space Administration (NASA) was the first to apply it to construct virtual space machine models. Existing DT works and implementations are still in their early stages, according to the literature, and require a lot of effort. However, it is well-integrated into a wide range of applications, such as biomedical systems, manufacturing, aerospace, agriculture, smart cities, and weather forecasting [1,2]. In addition, a number of distinct specialist engineers and computer scientists are needed in this essential subject to design an efficient DT system for any physical system. Their responsibilities will include building and designing the essential product prototype as well as creating a detailed explanation of the virtual system.

With the advance of modern technologies currently, the digital twin is integrating with other technologies like cloud computing, Artificial Intelligence (AI), Big data, Blockchain, and IoT and sensor data fusion to develop a digital data structure that is reorganized and modified when its physical counterparts alter [7,12,16–20]. In recent years, IoT integration with DT has allowed for greater visibility of physical Twins and their current status. It also simplifies the connection and documentation of the system function so that the performance of the physical devices may be understood and clarified. Furthermore, Artificial Intelligence is continually improving DT's abilities by processing data from its physical counterpart and surroundings. Data analysis and prediction result in the detection of patterns, data classification, and model identification.

The introduction of modern technologies, such as Blockchain, has recently made it easier to use Digital Twins (DTs) in practical applications [21–23]. Also, emerging AI technology like deep learning with CPSs is utilized and accompanied by digital twins to improve performance and enhance manufacturing production [17]. It proposes a concept for smart manufacturing and Industry 4.0 through the use of conceptual schema and modern technology. In [24], they presented a novel scheme for fault detection and prognosis based on digital twins and deep learning for fault diagnosis. The deep learning system may easily detect failure in the process. Digital Twins are used for defecting diagnosis for an electrical machine, which is proposed in [22]. A new DT with the Industrial Internet is introduced in [25] to improve the vision for smart production [26–28].

Thus, this research paper presents an efficient fault diagnosis framework based on GA and ML algorithms. The contributions of this paper are concise as follows:

- Developing a hybrid adaptive ML and GA algorithm for automatic and efficient fault diagnosis in industrial IoT applications.
- Developing a proficient remote monitoring system for early IIoT-based fault detection and diagnosis using DT concepts in intelligent industrial control systems.
- Achieving an experimental analysis of the proposed framework in classifying faults using simulated data with a lower cost for industrial IoT applications.
- Performing comprehensive comparisons that prove the high classification accuracy of faults diagnosis for the proposed framework compared to related frameworks.

The remainder of this work is structured as follows. Section 2 discusses the previous related works. All details and descriptions of the suggested framework are given in Section 3. Then, the high-level structure of the suggested framework with further details is presented in Section 4. The achieved outcomes for examining the performance of the suggested framework are offered in Section 5. Section 6 presents the main concluding remarks and some future directions.

## 2 Previous Works

This study aims to create and use a DT model for the Triplex pump by using simulated data to detect and diagnose faulty equipment. To increase efficiency, safety, and continuous production, fault detection and diagnosis have been carried out for various industrial processes. Several intelligent fault diagnosis technologies have been presented in recent decades to improve the durability and safety of complicated equipment [29,30]. The model-based approach, in which a precise model of complex equipment is created using various analytical expressions, is a well-known and successful method for fault diagnosis [31–34]. The well-established model-based approaches have been effectively deployed on numerous processes for complex equipment based on physical and mathematical knowledge of industrial processes. But developing an accurate model with a thorough understanding of the underlying physical processes is a problem because of the enormous complexity of some equipment.

Another powerful way to predict or diagnose fault states in complex equipment is the data-driven approach [35–38]. The mathematical correlations reflect the equipment's complicated features; without any physical model, the features of a fault condition can be learned. Recently, machine learning algorithms have gained great attention and are used in mechanical fault diagnosing operations. The use of a backpropagation neural network based on the Levenberg-Marquardt training algorithm to develop an effective fault detection strategy for proton exchange membrane fuel cells (PEMFCs) systems is presented in [39]. Another fault diagnosis based on a multiple kernel support vector machine for the distillation column is developed in [40]. Dynamic distillation column simulation generates data

in normal and faulty operations and then uses the proposed approach to provide a high fault detection rate. Deep learning based on LeNet-5 is proposed in [40] for fault diagnosis of motor, centrifugal, and hydraulic pumps. The Convolution Neural Network (CNN) is proposed in [41] for detecting Photovoltaics (PV) inverter open-circuit problems. Finally, the edge cloud paradigm is used to perform real-time diagnosis.

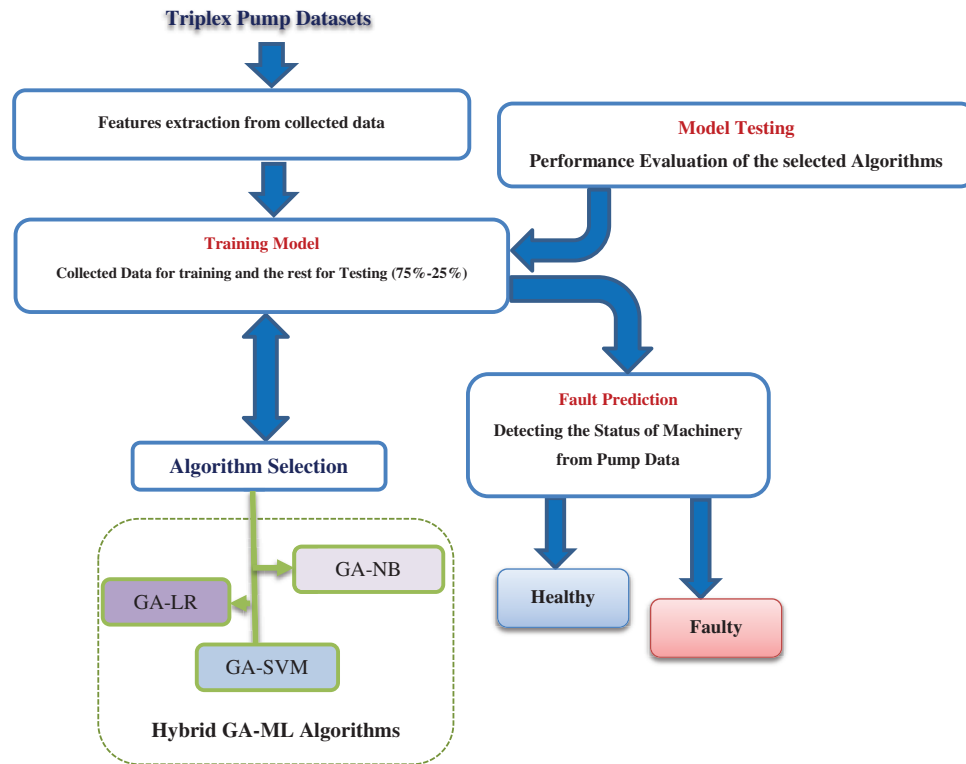
It is observed that work presented many works on faults diagnosis. However, there are rare contributions to the literature on digital twin-assisted machine learning algorithms in automated industrial IIoT applications. Consequently, this paper suggests an intelligent fault prediction framework that composes four steps precisely: (1) Data Acquisition Step (DAS), (2) Data Handling Step (DHS), (3) AI-based Model Training and Testing (AIMT2), and (4) Fault Classification and Detection Step (FCDS). This research proposes to build an innovative smart prediction digital twin-assisted AI framework using a hybrid genetic algorithm using various ML algorithms such as Logistic Regression (LR), Naïve Bayes (NB), and Support Vector Machine (SVM) that can effectively classify faults.

### 3 Suggested Automated Fault Diagnosis Framework

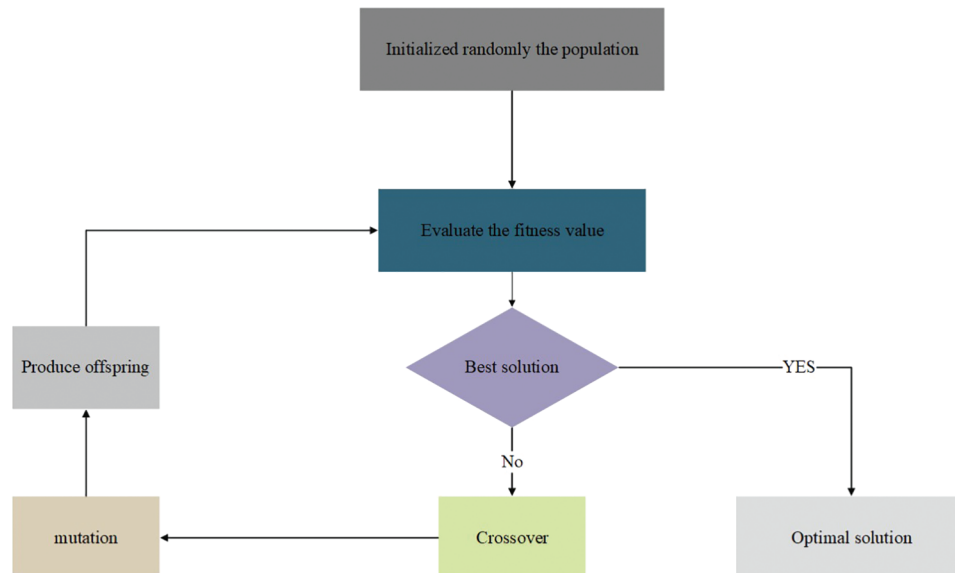
In this work, we presented a novel automated fault diagnosis framework for discovering the status of digital twin-assisted AI. Fig. 1 demonstrates the steps of the suggested framework. It composes different algorithms, such as GA, LR, NB, and SVM methods. Additionally, the suggested hybrid automated framework comprises various significant steps to achieve the diagnostic system, which are clarified as follows:

- **Step 1 (Data Gaining and Gathering):** This step is responsible for data collection from DT sources to be utilized in the next steps in the suggested framework.
- **Step 2 (Data Handling):** This step aims to accumulate all gathered data in the previous step and organize them in one dataset. This structured dataset is loaded to the subsequent steps to demonstrate the status of all collected faults.
- **Step 3 (Model Building):** The main goal of this step is to build the employed ML models. This can be achieved by firstly dividing the collected preprocessed data into 75% and 25% for training and testing processes, respectively. Next, the random subsampling selection process for the training data is applied. Finally, different assessment metrics are utilized to comprehensively evaluate the suggested framework.
- **Step 4 (Fault Classification and Detection):** In this phase, the suggested framework is tested by forwarding the testing data to the developed trained ML models to check the classification efficacy of each examined ML model.

The layout of the employed genetic algorithm used in the suggested DT-based AI framework is given in Fig. 2. Genetic algorithm (GA) is utilized to optimize the features of the model through minimizing the fitness function of the error between the predicted and the actual values of the measured data.



**Figure 1:** The suggested DT-based AI framework for fault diagnoses

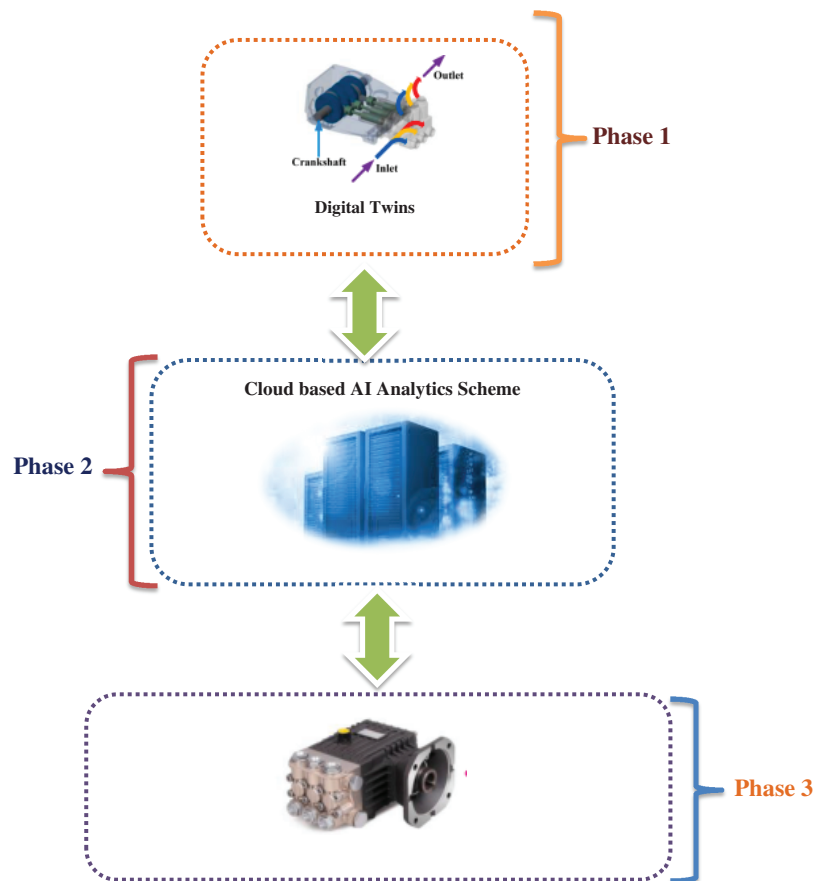


**Figure 2:** The steps of the genetic algorithm

#### 4 High-Level Structure of the Suggested DT-Based IIoT Framework

The suggested framework has the objective of professionally fault diagnosis for machinery. So, the diagnosis process is mainly accomplished by exploiting the DT, IIoT, CC, and AI concepts for real-time fault diagnosis and monitoring of industrial systems. Thus, this proposes enlightening the efficiency of industrial control systems by delivering a more advantageous intelligent system that enables early observation of the faulty process. The proposed framework consists of three key phases that cooperate to accomplish system purposes. Fig. 3 explains the suggested fault diagnosis framework that composes three different phases.

- **Phase 1:** In this phase, the IIoT system is utilized for real-time DT data gathering.
- **Phase 2:** In this phase, the cloud server is considered the main storage & processing pool for handling the generated data.
- **Phase 3:** The industrial operators use a cloud-based dashboard monitoring system to check machinery data from the industrial system. These industrial operators will finally make the proper decision based on the reports received from the cloud server.



**Figure 3:** Projected remote monitoring framework for early fault detection and diagnosis

## 5 Results and Analysis

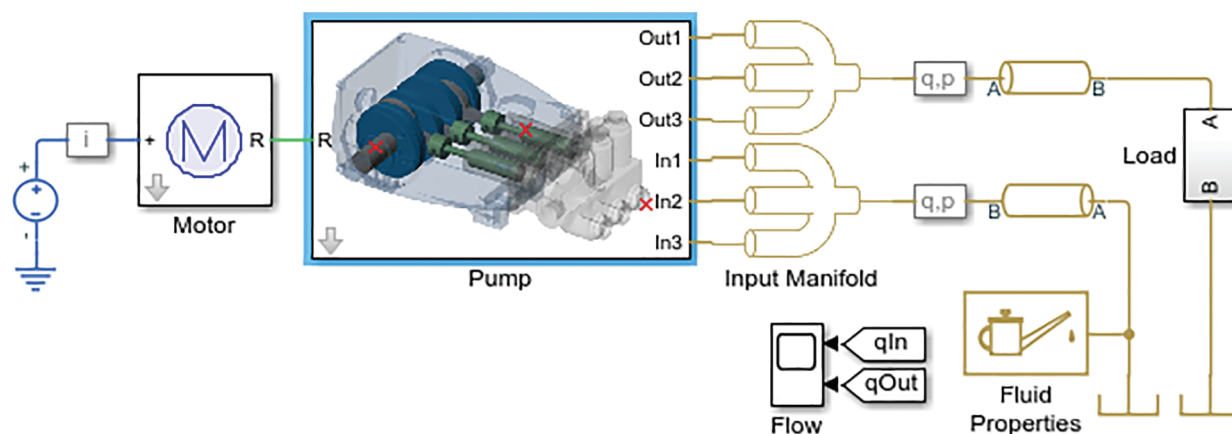
This section offers the simulation results, their analysis, and discussions of the suggested diagnosis framework.

### 5.1 Simulation Environment

The simulation environment that is utilized to produce the results for the suggested fault diagnosis framework is a laptop with i5 CPU Intel (R) Core (TM) and 12 GB RAM. All employed algorithms are developed using the Python programming language and carried out on Windows 10. These developed ML models are tested on the collected triplex pump data.

## 5.2 Triplex Pump Dataset

The simulated model of the Triplex pump is created using Simscape in MATLAB, as shown in Fig. 4. It is introduced to provide an automated fault diagnosis algorithm. Different 240 measurements of flow and pressure under different fault conditions are obtained generated using predictive model toolbox for the simulated model in the MATLAB Simulink [42] and put in the form of a data set. Then, we expand the data set by increasing the number of feature samples, improving classification model accuracy. Consequently, we extract time-domain features such as signal mean, root mean square, the peak value of signals standard deviation, signal-to-noise ratio, and Total harmonic distortions for pressure and flow signals. Likewise, analyzing the frequency spectrum for both the flow and the pressure will emphasize the cyclic nature of the flow signal and could give better insight into how the flow signal changes under different fault conditions. Afterward, computing spectral features such as peaks, modal coefficients, and band power.



**Figure 4:** The simulated model of the Triplex pump

### 5.3 Assessment Classification Metrics

Different assessment classification metrics are utilized to comprehensively evaluate the suggested fault diagnosis framework, such as F1-score, recall, precision, and accuracy. In addition, the confusion matrix [41] given in Table 1 is used to show the correct and incorrect classification percentages. In Table 1,  $P$  refers to positive classified samples, and  $N$  refers to the negative classified samples for the employed ML classifier.



**Table 1:** Confusion matrix

	Predicted healthy	Predicted faulty
Actual healthy	TP	FN
Actual faulty	FP	TN

The mathematical formulas for the employed assessment classification metrics are presented in Eqs. (1) to (4) [41].

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

where  $FP$  is a collection of regular instances classified as an anomaly scenario  $FP$ ,  $TN$  is an incorrectly measured number of regular instances, and  $FN$  is a list of anomalies observed as ordinary scenarios.

#### 5.4 Results Analysis

In the analysis of the suggested framework, the classification performance of each employed ML classifier is individually assessed, as demonstrated in the obtained results in Table 2. The utilized data has two statuses: healthy (i.e., Normal) or faulty status. So, the classification of the generated pump data is considered a binary classification process. In the simulation analysis, all data in the training process are randomly selected during the training process.

**Table 2:** Results of ML classifiers for fault detection based on simulated data as healthy or faulty status

Classifier	Diagnosis status	Precision	Recall	F1-score
LR	Healthy	0.69	0.82	0.75
	Faulty	0.96	0.92	0.94
NB	Healthy	0.61	1.00	0.76
	Faulty	1.00	0.86	0.92
SVM	Healthy	0.77	0.91	0.83
	Faulty	0.98	0.94	0.96



Tables 2 to 5 show the attained simulation results of all examined cases for the suggested framework. Fig. 5 shows the obtained confusion matrix, while Fig. 6 shows the receiver operating characteristic (ROC) curve for all tested algorithms in the suggested fault diagnosis framework. The experimental results demonstrated that the hybrid GA-ML method achieves outstanding results compared to using only ML methods, like LR, NB, and SVM. It is also observed that the suggested hybrid framework accomplishes a high detection accuracy of 95% when using the hybrid GA-SVM classifier.

**Table 3:** Results of suggested hybrid GA-ML classifiers for fault detection based on simulated data as healthy or faulty status

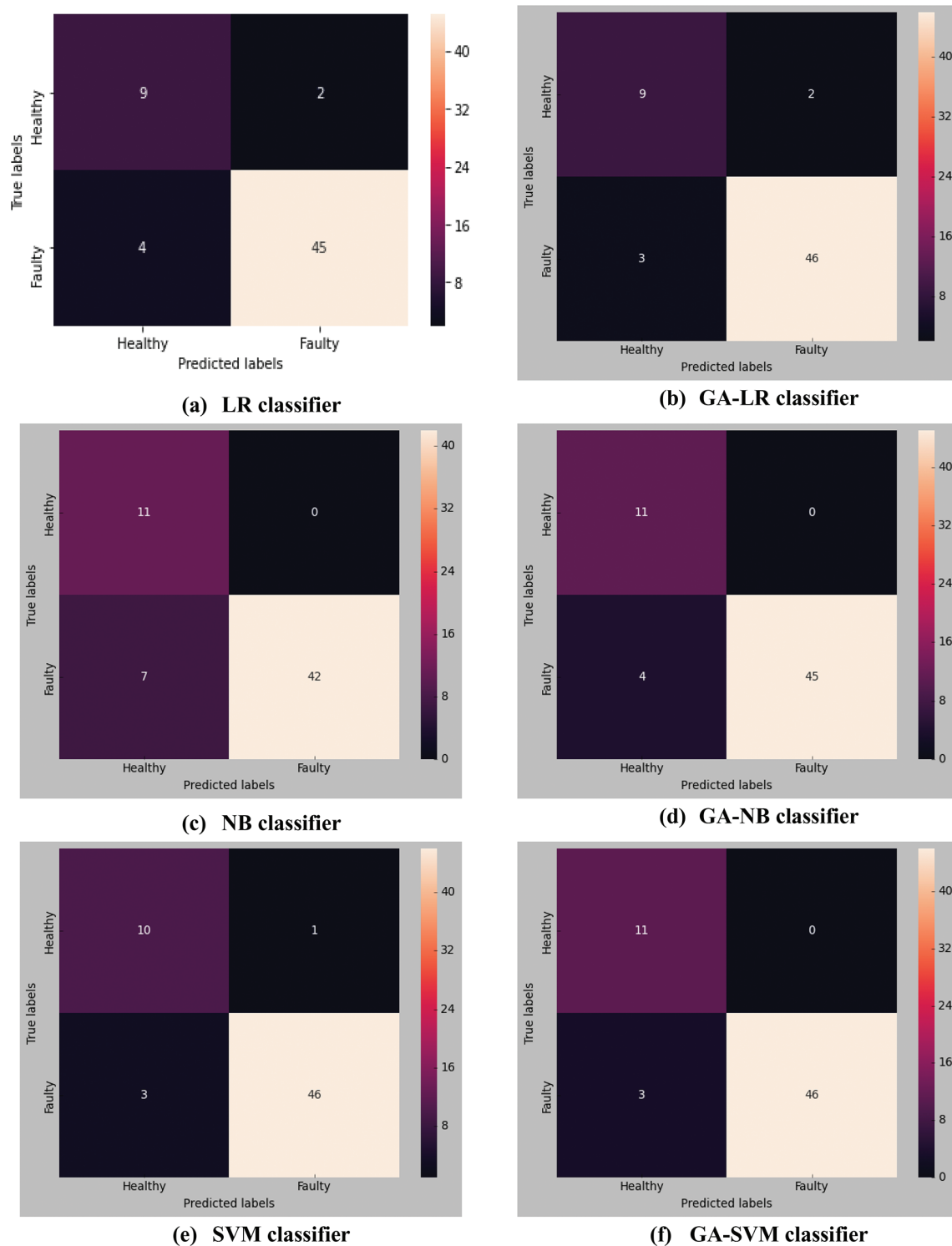
Classifier	Diagnosis status	Precision	Recall	F1-score
GA-LR	Healthy	0.75	0.82	0.78
	Faulty	0.96	0.94	0.95
GA-NB	Healthy	0.73	1.00	0.85
	Faulty	1.00	0.92	0.96
GA-SVM	Healthy	0.79	1.00	0.88
	Faulty	1.00	0.94	0.97

**Table 4:** Average results of ML classifiers for fault detection

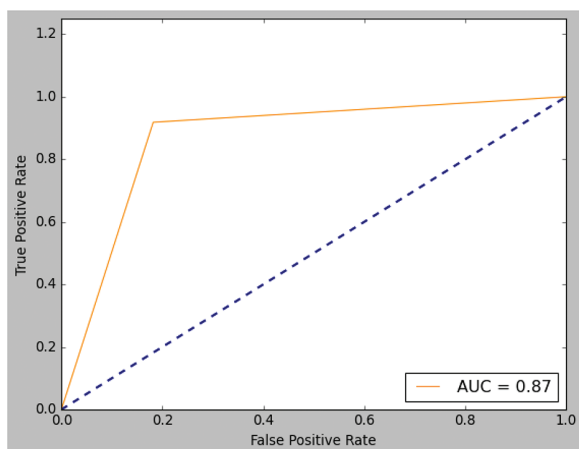
Classifier	Accuracy	Precision	Recall	F1-score
LR	0.90	0.82	0.87	0.84
NB	0.88	0.81	0.93	0.84
SVM	0.93	0.87	0.92	0.90

**Table 5:** Average results of suggested hybrid GA-ML classifiers for fault detection

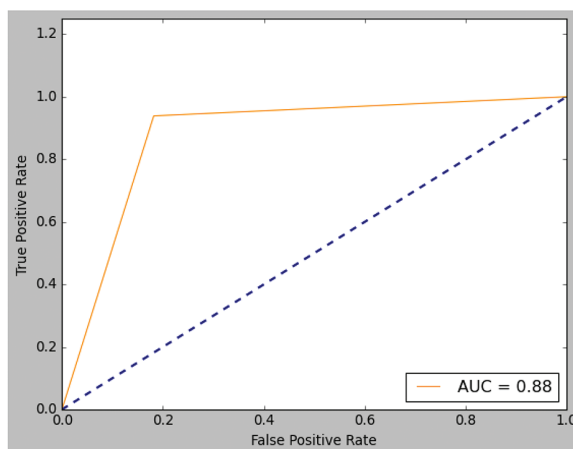
Classifier	Accuracy	Precision	Recall	F1-score
GA-LR	0.92	0.85	0.88	0.87
GA-NB	0.93	0.87	0.96	0.90
GA-SVM	0.95	0.89	0.97	0.92



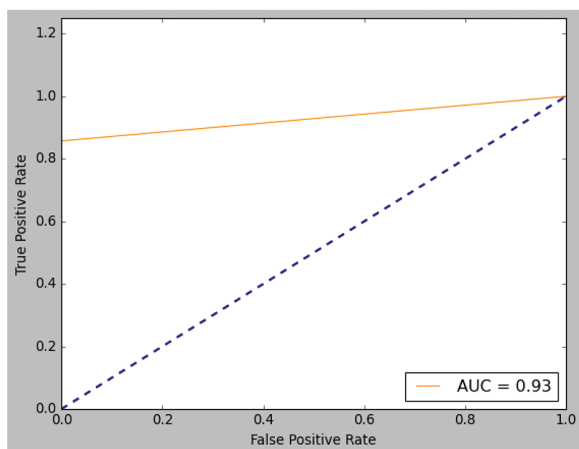
**Figure 5:** Confusion matrices of all examined classifiers in the suggested framework



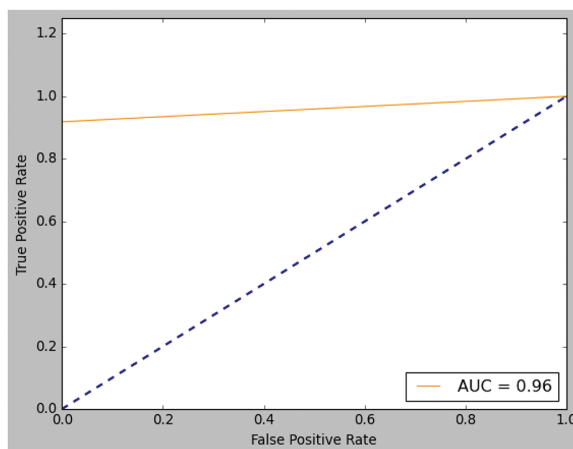
(a) LR classifier



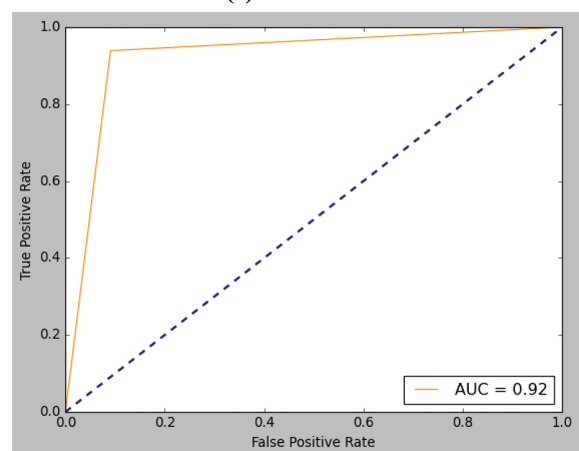
(b) GA-LR classifier



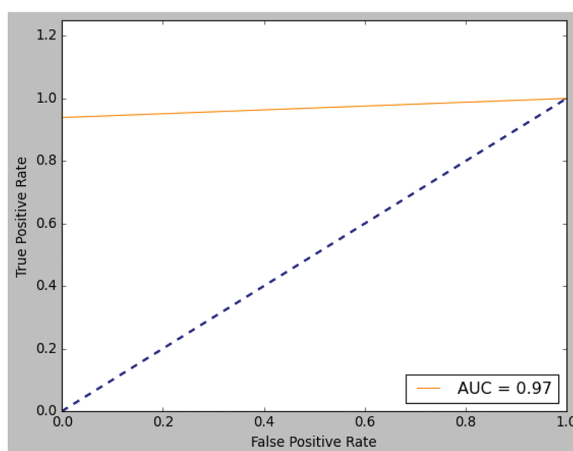
(c) NB classifier



(d) GA-NB classifier



(e) SVM classifier



(f) GA-SVM classifier

**Figure 6:** ROC curves of all examined classifiers in the suggested framework

## 6 Conclusions and Future Work

With the advancement in industry 4.0, many technologies such as Cloud Computing (CC), Digital Twins (DT), Artificial Intelligence (AI), and Industrial Internet of Things (IIoT) can be utilized to improve the performance of industrial systems. Therefore, this work presented an efficient fault diagnosis system utilizing the integration of DT and AI to accurately identify the early faults in industrial machines. In the suggested framework, the DT model of the industrial machine was constructed by generating the simulation testing model and further updating incessantly measured data from the physical asset. Then, different ML algorithms were employed in the suggested fault detection and diagnosis framework for data classification. The suggested hybrid framework was examined through a case study of triplex pump fault diagnosis with different generated datasets for learning purposes. The experimental results demonstrated that the hybrid GA-ML method elasticities outstanding results compared with using ML methods individually. It is observed that the suggested hybrid framework accomplished a high detection accuracy of 95% in the case of using the hybrid GA-SVM classifier. Therefore, the suggested fault diagnosis framework can assist industrial operators in making proper decisions regarding fault analysis for industrial control systems. In the future, different efficient deep learning algorithms must be employed in the suggested framework to further improve detection performance and diagnosis efficiency. In addition, different industrial IoT scenarios and applications can be investigated for developing other DT-based automated diagnosis systems.

**Acknowledgement:** Princess Nourah bint Abdulrahman University Researchers Supporting Project Number (PNURSP2022R197), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

**Funding Statement:** Princess Nourah bint Abdulrahman University Researchers Supporting Project Number (PNURSP2022R197), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

- [1] A. Rasheed, O. San and T. Kvamsdal, "Digital twin: Values, challenges and enablers from a modeling perspective," *IEEE Access*, vol. 8, no. 3, pp. 21980–22012, 2020.
- [2] D. Jones, C. Snider, A. Nassehi, J. Yon and B. Hicks, "Characterising the digital twin: A systematic literature review," *CIRP Journal of Manufacturing Science and Technology*, vol. 29, no. 2, pp. 36–52, 2020.
- [3] A. Fuller, Z. Fan, C. Day and C. Barlow, "Digital twin: Enabling technologies, challenges and open research," *IEEE Access*, vol. 8, no. 2, pp. 108952–108971, 2020.
- [4] G. Schroeder, C. Steinmetz, R. Rodrigues, R. Henriques and C. Pereira, "A methodology for digital twin modeling and deployment for industry 4.0," *Proceedings of the IEEE*, vol. 109, no. 4, pp. 556–567, 2020.
- [5] F. Tao, H. Zhang, A. Liu and A. Nee, "Digital twin in industry: State-of-the-art," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2405–2415, 2019.
- [6] Q. Qi and F. Tao, "Digital twin and big data towards smart manufacturing and industry 4.0: 360-degree comparison," *IEEE Access*, vol. 6, no. 2, pp. 3585–3593, 2018.
- [7] C. Wagner, J. Grothoff, U. Eppe, R. Drath and P. Zimmermann, "The role of the industry 4.0 asset administration shell and the digital twin during the life cycle of a plant," in *Proc. of 22nd IEEE Int. Conf. on Emerging Technologies and Factory Automation (ETFA)*, Limassol, Cyprus, pp. 1–8, 2017.

- [8] Q. Qi and F. Tao, "Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison," *IEEE Access*, vol. 6, no. 5, pp. 3585–3593, 2018.
- [9] W. Kritzing, M. Karner, G. Traar, J. Henjes and W. Sihn, "Digital twin in manufacturing: A categorical literature review and classification," *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 1016–1022, 2018.
- [10] S. Boschert and R. Rosen, "Digital twin-the simulation aspect," *International Journal of Mechatronic Futures*, vol. 3, no. 5, pp. 59–74, 2016.
- [11] A. El-Saddik, "Digital twins: The convergence of multimedia technologies," *IEEE Multimedia*, vol. 25, no. 2, pp. 87–92, 2018.
- [12] K. Alam and A. El-Saddik, "C2PS: A digital twin architecture reference model for the cloud-based cyber-physical systems," *IEEE Access*, vol. 5, no. 3, pp. 2050–2062, 2017.
- [13] B. Schleich, N. Anwer, L. Mathieu and S. Wartzack, "Shaping the digital twin for design and production engineering," *CIRP Annals of Manufacturing Technology*, vol. 66, no. 1, pp. 141–144, 2017.
- [14] M. Grieves and J. Vickers, "Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems," *Transdisciplinary Perspectives on Complex Systems*, vol. 4, no. 12, pp. 85–113, 2017.
- [15] Y. Jiang, S. Yin, K. Li, H. Luo and O. Kaynak, "Industrial applications of digital twins," *Philosophical Transactions of the Royal Society*, vol. 3, no. 7, pp. 2207–2219, 2021.
- [16] S. Huang, G. Wang, Y. Yan and X. Fang, "Blockchain-based data management for digital twin of product," *Journal of Manufacturing Systems*, vol. 54, no. 7, pp. 361–371, 2020.
- [17] J. Lee, M. Azamfar, J. Singh and S. Siahpour, "Integration of digital twin and deep learning in cyber-physical systems: Towards smart manufacturing," *IET Collaborative Intelligent Manufacturing*, vol. 2, no. 1, pp. 34–36, 2020.
- [18] F. Tao, J. Cheng, Q. Qi, M. Zhang and F. Sui, "Digital twin-driven product design, manufacturing and service with big data," *The International Journal of Advanced Manufacturing Technology*, vol. 94, no. 9, pp. 3563–3576, 2018.
- [19] Y. Cai, B. Starly, P. Cohen and Y. Lee, "Sensor data and information fusion to construct digital-twins virtual machine tools for cyber-physical manufacturing," *Procedia Manufacturing*, vol. 10, no. 3, pp. 1031–1042, 2017.
- [20] T. Heikkinen, "Case study: Challenges of transition from a manufacturing company to a maintenance partner," *International Journal of Service Science, Management, Engineering, and Technology*, vol. 3, no. 4, pp. 1–12, 2012.
- [21] J. Wang, L. Ye, R. Gao, X. Li and L. Zhang, "Digital twin for rotating machinery fault diagnosis in smart manufacturing," *International Journal of Production Research*, vol. 57, no. 12, pp. 3920–3934, 2019.
- [22] E. Hinchy, N. Dowd and C. McCarthy, "Using open-source microcontrollers to enable digital twin communication for smart manufacturing," *Procedia Manufacturing*, vol. 3, no. 8, pp. 1213–1219, 2019.
- [23] F. Tao, F. Sui, A. Liu, Q. Qi and A. Nee, "Digital twin-driven product design framework," *International Journal of Production Research*, vol. 57, no. 12, pp. 3935–3953, 2019.
- [24] W. Booyse, D. Wilke and S. Heyns, "Deep digital twins for detection, diagnostics and prognostics," *Mechanical Systems and Signal Processing*, vol. 14, no. 10, pp. 106–121, 2020.
- [25] J. Cheng, H. Zhang, F. Tao and C. Juang, "DT-II: Digital twin enhanced industrial internet reference framework towards smart manufacturing," *Robotics and Computer-Integrated Manufacturing*, vol. 6, no. 2, pp. 101–108, 2020.
- [26] K. Randhawa, C. Loo, M. Seera, C. Lim and A. Nandi, "Credit card fraud detection using AdaBoost and majority voting," *IEEE Access*, vol. 6, no. 6, pp. 14277–14284, 2018.
- [27] S. Forrest, "Genetic algorithms," *ACM Computing Surveys*, vol. 28, no. 1, pp. 77–80, 1960.
- [28] R. Qu, E. Burke, B. McCollum, L. Merlot and Y. Lee, "A survey of search methodologies and automated system development for examination timetabling," *Journal of Scheduling*, vol. 12, no. 1, pp. 55–89, 2009.
- [29] T. Zhang, J. Chen, F. Li, K. Zhang and E. Xu, "Intelligent fault diagnosis of machines with small & imbalanced data: A state-of-the-art review and possible extensions," *ISA Transactions*, vol. 11, no. 9, pp. 152–171, 2022.

- [30] Y. Cheng, M. Lin, J. Wu, H. Zhu and X. Shao, "Intelligent fault diagnosis of rotating machinery based on continuous wavelet transform-local binary convolutional neural network," *Knowledge-Based Systems*, vol. 21, no. 6, pp. 132–145, 2021.
- [31] X. Jin, R. RobertJeremiah, T. Su and J. Kong, "The new trend of state estimation: From model-driven to hybrid-driven methods," *Sensors*, vol. 21, no. 6, pp. 2085–2099, 2021.
- [32] Z. Gao, C. Cecati and S. Ding, "A survey of fault diagnosis and fault-tolerant techniques-part I: Fault diagnosis with model-based and signal-based approaches," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 6, pp. 3757–3767, 2015.
- [33] M. Zhong, T. Xue and S. Ding, "A survey on model-based fault diagnosis for linear discrete time-varying systems," *Neurocomputing*, vol. 30, no. 6, pp. 51–60, 2018.
- [34] C. Wang and Z. Xu, "An intelligent fault diagnosis model based on deep neural network for few-shot fault diagnosis," *Neurocomputing*, vol. 45, no. 6, pp. 550–562, 2021.
- [35] K. Tidiri, N. Chatti, S. Verron and T. Tiplica, "Bridging data-driven and model-based approaches for process fault diagnosis and health monitoring: A review of researches and future challenges," *Annual Reviews in Control*, vol. 4, no. 2, pp. 63–81, 2016.
- [36] A. Arbaoui, A. Ouahabi, S. Jacques and M. Hamiane, "Concrete cracks detection and monitoring using deep learning-based multiresolution analysis," *Electronics*, vol. 10, no. 15, pp. 17–30, 2021.
- [37] A. Samanta, S. Chowdhuri and S. Williamson, "Machine learning-based data-driven fault detection/diagnosis of lithium-ion battery: A critical review," *Electronics*, vol. 10, no. 11, pp. 130–139, 2021.
- [38] Y. Xing, B. Wang, Z. Gong, Z. Hou and K. Jiao, "Data-driven fault diagnosis for pem fuel cell system using sensor pre-selection method and artificial neural network model," *IEEE Transactions on Energy Conversion*, vol. 3, no. 5, pp. 1–19, 2022.
- [39] S. Taqvi, H. Zabiri, F. Uddin, M. Naqvi and A. Maulud, "Simultaneous fault diagnosis based on multiple kernel support vector machine in nonlinear dynamic distillation column," *Energy Science & Engineering*, vol. 5, no. 3, pp. 1–17, 2022.
- [40] X. Wang, B. Yang, Z. Wang and X. Guan, "A compressed sensing and CNN-based method for fault diagnosis of photovoltaic inverters in edge computing scenarios," *IET Renewable Power Generation*, vol. 2, no. 7, pp. 1–12, 2022.
- [41] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information Processing & Management*, vol. 45, no. 4, pp. 427–437, 2009.
- [42] L. Rivera, M. Scholz, M. Fritscher and K. Schilling, "Towards a predictive maintenance system of a hydraulic pump," *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 447–452, 2018.