Enhancing Production System Performance: Failure Detection and Availability Improvement with Deep Learning and Genetic Algorithm

Artika Farhana[1], Shaista Sabeer[2], Ayasha Siddiqua[3], Afsana Anjum[4]
Lecturer, Department of Computer Science, Jazan University, Saudi Arabia[1]
Lecturer, Department of Computer Science & Information Technology, Jazan University, Saudi Arabia[2, 3]
Lecturer, Department of Computer Science & Information Technology, Jazan University Jazan, KSA[4]

Abstract*—A crucial component of industrial operations is the detection of production system failures, which aims to spot any problems before they get worse. By applying cutting-edge methods like deep learning and genetic algorithms, failure detection accuracy may be improved, allowing for preemptive actions to reduce downtime and maximize system availability. These methods improve reactivity to possible errors and solve dynamic issues, which enhances the overall efficiency and reliability of production systems. This study offers a novel method for improving the availability and failure detection of production systems using deep learning techniques and genetic algorithms in a data-driven strategy. The goal of the project is to provide a complete framework for efficient failure detection that incorporates deep learning models, particularly Convolutional Neural Network (CNN) Autoencoder. Furthermore, system configurations are optimized through the use of genetic algorithms, improving overall availability. The suggested model is able to identify complex patterns and connections in the data by being trained on a variety of datasets that contain information about equipment failure. The incorporation of genetic algorithm guarantees flexibility and resilience in system setups, hence augmenting total availability. The study presents a proactive and flexible approach to the dynamic issues encountered in industrial environments, providing a notable breakthrough in failure detection and availability improvement. The proposed model is implemented in Python software. It achieves an astounding 99.32% accuracy rate, which is 3.58% higher than that of current techniques like CNN-LSTM (Long Short-Term Memory), Bi-LSTM (Bi-directional Long Short-Term Memory), and CNN-RNN (Recurrent Neural Network). The data-driven approach's high accuracy highlights its efficacy in forecasting and avoiding problems, which minimizes downtime and maximizes production efficiency.

***Keywords*—Autoencoder; availability enhancement; convolutional neural network; failure detection; genetic algorithm**

I. INTRODUCTION

Improving the performance of production systems on an ongoing basis is a primary goal in the manufacturing and industrial domains [1]. It is a result of the necessity to maximize productivity, reduce downtime, and guarantee the reliable provision of high-quality goods and services. Production systems must adapt to the competitive environment by moving beyond conventional paradigms and embracing novel approaches in order to remain robust in the face of shifting consumer expectations and technology breakthroughs [2]. This study explores the crucial area of production system improvement in an effort to push the envelope by presenting cutting-edge tactics that go beyond accepted constraints [3]. It begins with the planning and design of the product, moves through the complex manufacturing procedures, and concludes with the distribution and pleasure of the consumer [4].

The intrinsic difficulties that production systems encounter, such as the requirement to strike a balance between competing demands like cost-effectiveness, quality control, and on-time delivery, necessitate improvement [5]. Reactive methods of system optimization, which deal with problems as they arise, frequently lead to inefficient use of resources and more downtime [6]. The emphasis is shifting towards proactive and predictive techniques that foresee problems before they become serious ones as industries seek for leaner, more flexible production processes. In order to improve production system performance, this research will critically assess current approaches and offer novel solutions that make use of state-of-the-art technology [7]. It attempts to close the knowledge gap between theoretical developments and real-world application by offering sectors looking to stay competitive in a time of quickening technology change and shifting consumer expectations practical insights [8].

The early detection and mitigation of possible system failures in production systems represent key problems in the context of industrial operations and critical infrastructure [9]. Conventional failure detection techniques can result in expensive downtime and inefficiencies since they are reactive in nature and frequently rely on rule-based systems or simple algorithms. Recognizing this, the current research offers a deep learning-based data-driven approach for failure detection, which represents a revolutionary paradigm change [10]. Using deep learning to uncover complex patterns from large datasets, this methodology deviates from traditional methods and improves the system's capacity to anticipate and proactively handle failure situations. [11].

Deep learning has the potential to improve failure detection accuracy and provide the necessary scalability for the complicated production situations of today [12]. Improving availability is a complex task that goes beyond maintaining system uptime. It entails striking a tactical balance between the effectiveness of maintenance, system dependability, and flexibility in response to shifting operating circumstances [13]. In order to identify optimal solutions within the large parameter space of a production system, genetic algorithms, as an optimization tool, replicate the process of evolution and provide a possible path towards overcoming these obstacles.

Conventional methods of accomplishing these goals frequently depend on reactive tactics, which deal with problems only after they occur, resulting in downtime and inefficient use of resources [14]. Genetic algorithms are incorporated into the framework as a complement to the deep learning paradigm in order to optimize system availability. Natural selection and evolution serve as the inspiration for genetic algorithms, which offer a potent way to discover and develop the best possible configurations for the production system. These algorithms constantly alter parameters, searching for configurations that minimize failure risk and maximize total system availability through recurrent processes of crossover, mutation, and selection. This study’s combination of genetic algorithms and deep learning offers a clever and comprehensive method for managing production systems. Through the integration of deep learning’s predictive powers with genetic algorithms’ optimization skills, the suggested framework seeks to create a proactive system that can both detect possible malfunctions and modify the system configuration to improve overall availability. The study is at the vanguard of efforts to transform the monitoring, maintenance, and optimization of production systems due to the synergy between innovative data-driven technologies.

Thekey contributions of the article is,

- The CNN Autoencoder based on deep learning is incorporated for failure detection. By automatically extracting pertinent features from the input data, the CNN Autoencoder improves the model's ability to identify complex patterns that may indicate probable problems, offering a reliable and data-driven method of failure detection.
- The Genetic Algorithm is employed to improve availability by optimizing system configurations in response to changing operating circumstances. This genetic algorithm-based method guarantees flexibility and resilience, tackling the intricacies of industrial production processes and leading to increased availability overall.
- The integration of CNN Autoencoder and Genetic Algorithm, results in a comprehensive approach for availability improvement and failure detection. By using cutting-edge data-driven methodologies to systematically handle both failure detection and system optimization elements, this all-encompassing strategy improves the overall resilience and efficiency of production systems.

There are five primary sections of the paper: In Section II, relevant works in the subject of production system failure detection are reviewed and current techniques are summarized. The issue statement is defined in Section III, along with the shortcomings and difficulties of the existing methods. The suggested technique, which combines genetic algorithms and deep learning to provide an advanced data-driven solution, is described in Section IV. The model's output is shown in Section V, along with a discussion of the model's performance and accuracy in relation to other techniques. The report is finally concluded in Section VI, which summarizes the main conclusions and offers directions for further research into the field of production system failure detection.

II. RELATED WORKS

In light of climate change and sustainability issues, this study discusses the pressing need for smart energy production and highlights the shortcomings of conventional first-principle model-based strategies in an environment of growing system scale and uncertainty [15]. The article provides a thorough analysis and emphasizes the ways in which Data-Driven Control (DDC) and Machine Learning (ML) approaches are used to the tracking, regulating, optimizing, and fault-detection of power production facilities. It provides a thorough analysis of the ways in which these cutting-edge techniques help to resolve ambiguities and improve the efficiency of both traditional thermal power generation and renewable energy sources. In regards to visibility, maneuverability, adaptability, economic viability, and safety, the article lists the benefits of ML and DDC. It is crucial to remember that the study does not go into great detail about the particular difficulties or restrictions related to the application of ML and DDC procedures to power production structures, leaving space for additional investigations to investigate the possible downsides and improve these strategies for real-world use. The inherent drawback of data-driven modeling becomes apparent when confronted with the constraints of both online and offline data, including issues such as data incompleteness stemming from loss, uncertainty, and bias.

In the framework of ring spinning technological advances, the study presents a unique method to preventative care with an emphasis on forecasting and health monitoring [16]. It suggests utilizing predictive analytics to create a data-driven preventive service system built on a regularized Deep Neural network (DNN). To keep an eye on crucial parts, the method makes use of a system of sensors that is built into the frames of spinning machines, each of which has many spindles. A GA is developed for multi-sensor assessments and prediction, demonstrating its efficiency in leveraging bigger amounts of data with comparatively small training data sets. With the use of a neural sensor network, the framework provides condition-based evaluation for every component in order to anticipate anomalies, disruptions, and failures in real time. The study does not, however, go into great detail on the shortcomings or difficulties of the suggested model, providing opportunity for more study to investigate such limits and improve the execution for more widespread industrial uses.

In order to forecast failure behavior in the industrial environment, the study investigates the use of sophisticated data analytics, more especially a data-driven Failure Mode and Effective Analysis (FMEA) approach [17]. Utilizing operational and historical data from investment in industrial items' usage stage, the technique applies DL models to improve maintenance scheduling visibility and decision assistance. A real-world scenario in the aviation industry is used to verify the structure, and the results show an astounding 95% accuracy in defect prediction. By incorporating these findings into a data-driven FMEA framework, the dependency on employee knowledge and skill is lessened, and variability in risk and failure probability predictions is eliminated. Nevertheless, the study does not go into great detail about any drawbacks or difficulties with the suggested technique, providing opportunity for more research into real-world implementation problems and wider application across other industrial settings.

This study explores the intricacies of semiconductor production, which is a multistep process that uses a variety of equipment's and sub processes to create miniaturized electronic circuits [18]. With the goal of improving the process of producing semiconductors, the study uses data mining, SPC, and data-driven decision-making frameworks to analyze production process information in great detail. The goal is to use the newest technologies powered by data to increase productivity and reliability. The study highlights the potential for major process enhancements and offers a thorough analysis of current procedures; however, it does not go into great length on the drawbacks or difficulties that come with using methods that use data in the production of semiconductors. Some issues are not completely addressed, such as possible biases, execution difficulties, and the adaptation of these approaches to varied production contexts. To evaluate the applicability and constraints of these methods based on data in a wider range of semiconductor production environments, more investigation is necessary.

This study tackles the critical problem of estimating RUL for equipment predictions, highlighting the importance of precise forecasts in reducing maintenance expenses and improving system dependability[19]. This study presents a new data-driven methodology called Convolutional Neural Network- Long-Short Term Memory- Particle Swarm Optimization (CNN–LSTM–PSO) hybrid DNN, which integrates conventional neural networks, LSTM networks, and CNN. In order to increase the accuracy of Remaining Useful Life (RUL) forecasting, this hybrid model seeks to identify spatial relations from time series with multiple variables data and retain nonlinear properties. To improve the efficiency of the network, the research uses PSO to optimize the network's hyper parameters. The suggested CNN–LSTM–PSO model is noteworthy for its ability to provide multi-step-ahead recommendations. Utilizing a NASA-provided lithium-ion battery dataset, the validation test shows that the CNN–LSTM–PSO model outperforms other cutting-edge ML and Deep Learning (DL) techniques when evaluating a variety of efficiency metrics. To allow for more research into the suggested model's relevance across various datasets and commercial situations, the report nevertheless fails to go into great detail about its drawbacks or possible drawbacks. Further investigation might examine the model's applicability to other machinery kinds and operating environments.

This paper explores the topic of smart production, with a particular emphasis on utilizing Artificial Intelligence (AI), ML, and sophisticated data analysis to optimize semiconductors production processes [20]. It emphasizes how Industrial Internet of Things (IIOT) sensors are being used in production more and more, which means effective data management is required. To solve issues in semiconductor production, the study suggests a dynamic method that combines algorithms for neural networks with evolutionary programming. In particular, the study presents a novel feature selection technique employing neural networks and evolutionary algorithms to improve the production process efficiency. Although the research offers a detailed analysis and innovative approaches, it does not fully address any potential drawbacks or difficulties related to the suggested dynamic algorithms and smart feature selection. Additional investigation is necessary due to practical issues, the algorithms' adaptation to a variety of industrial contexts, and possible limitations. Subsequent investigations have to evaluate the resilience and constraints of these suggested remedies in diverse semiconductor production environments and take into account practical implementation difficulties.

In order to anticipate the initial yields in semiconductor production, the research presents a combined structure that focuses on finding the best companion combinations for the Final Testing (FT) procedure [21]. In order to create an efficient prediction of yield model, the process entails converting categorical information into multivariate vectors using the entity anchoring technique and assessing several ML methods. A Genetic Algorithm (GA) incorporated in the yield prediction framework is used to find the optimal accessory combinations, optimizing for the greatest initial yield estimation, after the best-fit ML model has been identified. By combining the strengths of ML and GA, a smart prediction method is created that stabilizes the Overall Equipment Effectiveness (OEE) for the FT process and reduces the negative effects of improper accessory pairings on yield rates. But the report does not go into great detail about any drawbacks or difficulties with the proposed design, thus there is need for more research to examine concerns of adaptability, versatility, and practical application in various semiconductor production settings. To evaluate the combined framework's resilience and generality in various production contexts, more research is necessary.

In addition to highlighting reliability's critical role in modern production facilities, the article also addresses reliability's influence on systems lives, expenses for upkeep, and repair charges [22]. Although a number of reliability modelling approaches have been investigated, including Fault Trees, Petri Nets, and Markov Chains, the process of developing dependability models is still demanding of labor and dependent on expertise. The report suggests using data from contemporary manufacturing facilities for automation or assist in the creation of dependability models as a solution to this problem. With an emphasis on information-driven reliability evaluation for cyber-physical machines, the suggested methodology seeks to capitalize on the abundance of data produced in sophisticated production lines. A case study is included to test and improve the suggested data-driven strategy from a practical standpoint. Nevertheless, the study does not go into great detail about the possible drawbacks or difficulties with the suggested structure, providing opportunity for further investigation into real-world application problems, potential exaggerations in the data, and the applicability of the method in various manufacturing contexts. To evaluate the security and sustainability of the data-driven resilience evaluation in actual manufacturing environments, more research is necessary.

In order to increase system accessibility and lower life cycle costs, the study tackles the crucial job of forecasting the RUL of systems [23]. Using numerous sensor time series indications, a Deep Long-Short Term Memory (DLSTM) network-based technique is introduced in this suggested solution. Through the usage of its DL framework, the DLSTM model is intended to fuse the aforementioned signals in order to provide precise predictions for RUL by revealing latent long-term relationships. The DLSTM's attributes and network layout are effectively tuned for accurate and reliable predictions in this article using an adaptive moment assessment approach and a grid-based searching strategy. Utilizing two turbofan engine datasets, the DLSTM model's efficacy is verified, showing favorable outcomes when contrasted to other neural network algorithms and the latest methods documented in the available literature. Though there is room for additional studies to examine real-world difficulties in implementation, possible prejudices in the data, and the framework's flexibility to various system forms and operational circumstances, the article does not go into great detail about potential drawbacks or difficulties connected with the recommended DLSTM model. To evaluate the DLSTM model's adaptability and generalization in actual manufacturing environments, more research is necessary.

The essential subject of forecasting RUL in diverse engineering and manufacturing scenarios is the focus of the literature review. The use of sophisticated deep learning and machine learning methods, such LSTM networks, is a recurring topic in the development of precise and effective RUL forecasts. These models make use of many sensor time series signals, which makes it possible to identify complex patterns and hidden connections in the data. The importance of RUL prediction in raising system availability, cutting life cycle costs, and improving maintenance plans is emphasized in the studies. To ensure optimal performance, model parameters are often tuned using grid search techniques and adaptive algorithms. Experimental validation on various datasets, such as turbofan engines, consistently shows these models to perform as robustly and competitively against other neural network designs and state-of-the-art methodologies. All of these studies have one thing in common, though: they refrain from going into great detail about the difficulties, biases, and real-world implementation problems that come with using these sophisticated predictive models in industrial settings. This leaves space for future research to address these important issues.

III. Problem Statement

The current problem in industrial environments is the inefficiencies and disruptions brought about by unanticipated breakdowns in production systems, which result in more downtime, less efficient use of resources, and weakened system availability overall. These problems are not fully addressed by traditional reactive techniques to failure detection and system optimization. In order to address this issue, this research offers a novel solution: a deep learning-based, data-driven method for failure detection and availability augmentation that makes use of genetic algorithms. The challenge at hand is creating a comprehensive approach that uses genetic algorithms to dynamically optimize system configurations for increased availability and deep learning techniques to proactively identify failure antecedents in production systems. The goal is to revolutionize the way that production system management is now done by offering a proactive, adaptable framework that not only foresees problems before they happen but also continuously improves the system to maximize availability and overall performance [24]. Because of its ability to improve pattern recognition in industrial data—especially in identifying minor symptoms of equipment failure—the suggested CNN Autoencoder-GA approach has been chosen. The utilization of Convolutional Neural Network (CNN) Autoencoder enables efficient recognition of intricate patterns, and the incorporation of Genetic Algorithms (GA) guarantees the adaptability and durability of system configurations in ever-changing industrial settings. By training on a variety of datasets, the data-driven approach improves the model's resilience and adaptability. On the other hand, the shortcomings of current approaches, like insufficient pattern recognition, static configurations, and insufficient forecasting abilities, make them less appropriate for handling the dynamic issues associated with effective failure detection and availability enhancement in operational systems.

IV. Proposed CNN Autoencoder – Genetic Algorithm Framework

In this paper, a unique data-driven technique for improving production system availability and failure detection is presented. The suggested system uses state-of-the-art techniques, such as deep learning and evolutionary algorithms, and integrates CNN Autoencoder for accurate failure detection. System configurations are optimized via genetic algorithms, increasing overall availability. The model, which has been trained on a variety of datasets, recognizes intricate patterns, and evolutionary methods guarantee system configuration flexibility. The model, which is implemented in Python, emphasizes how effective the data-driven approach is at issue prediction and prevention, downtime minimization, and production efficiency maximization. It is depicted in Fig. 1.

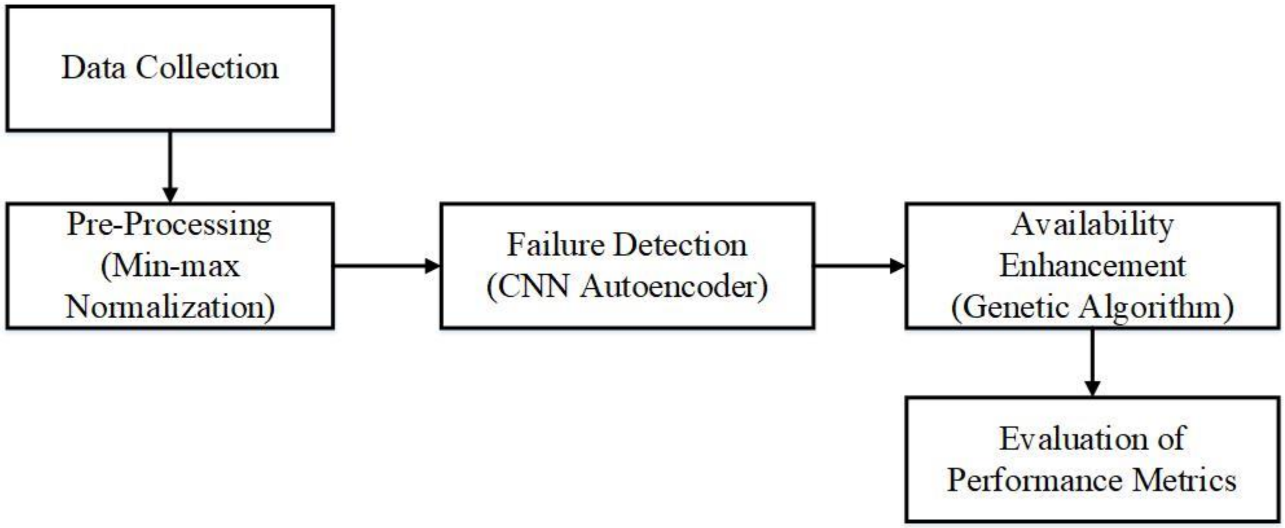

Fig. 1. Proposed methodology.

A. Data Collection

Kaggle, a well-known website for data science and machine learning contests, provided datasets for the equipment failure prediction. The meticulously selected datasets, which are publicly accessible on Kaggle, are an invaluable tool for practitioners and academics who are trying to create predictive models that may detect any malfunctions in machinery. The datasets available in Kaggle's repository cover a wide range of sectors and machinery kinds, making it possible to investigate various failure patterns and advance the creation of reliable prediction algorithms. These datasets are a great source of data for developing and accessing machine learning models, and they often contain elements like timestamps, operating parameters, and sensor readings. Researchers may test their models against pre-existing datasets by utilizing Kaggle's equipment failure prediction dataset platform. This approach promotes cooperation and creativity in the domain of predictive maintenance and reliability engineering [25].

B. Preprocessing using Min-Max Normalization

Preprocessing is essential when attempting to improve production system performance using a deep learning-based, data-driven strategy for genetic algorithm-based failure diagnosis and availability augmentation. In particular, applying Min-Max Normalization sticks out as an essential step in bringing the input data into compliance. This method makes sure that every variable contributes equally to the learning process by scaling the feature values within a given range, usually between 0 and 1. Normalizing the input data makes the deep learning model less susceptible to changes in the magnitude of various characteristics and more resilient, which supports steady and efficient training. By reducing the effect of different scales among input characteristics, Min-Max Normalization enhances the overall dependability of the model. This is important when it comes to failure detection and availability enhancement in intricate production systems. Moreover, the use of Min-Max Normalization is consistent with the overall objective of maximizing the accuracy and speed of convergence of the deep learning model. A more effective learning process is made possible by the standardized data distribution, which also helps to avoid particular traits from predominating during the training phase and thereby distorting the model's predictions. When it comes to availability enhancement and failure detection, wherein subtle patterns may signal approaching problems, and when accuracy is critical, the preprocessing step of Min-Max Normalization guarantees that the deep learning model is capable of identifying pertinent patterns and trends, which in turn enhances the model's ability to improve production system performance as shown in Eq. (1).

$$X_{Normalized} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

C. Deep Learning-based CNN Autoencoder for Failure Detection

In the field of industrial system failure detection, this study presents a new method based on Deep Learning using CNN Autoencoder architecture. The main goal is to create a complicated model that can learn nuanced patterns from large, complex information in order to proactively anticipate probable problems. A change from conventional techniques is marked by the use of CNN Autoencoders, which use neural networks' capacity to automatically extract pertinent characteristics from input data. The CNN design is very useful for jobs that need spatial connections, which makes it suitable for examining images, sensor data, and other multidimensional data sources that are frequently found in industrial settings.

An encoding phase that compresses input data into a latent representation and a decoding phase that reconstructs the input from this compressed representation are the fundamental workings of the CNN Autoencoder. The model gains the ability to reflect typical operating patterns throughout training, which makes it sensitive to variations that might signal impending problems. Deep learning guarantees a data-driven and adaptable approach to failure detection, able to identify intricate patterns that may defy conventional rule-based or heuristic techniques. Using this cutting-edge method, the research hopes to enhance the creation of reliable and effective failure detection systems that may improve the performance and dependability of industrial production systems.

The encoder, which is composed with a number of layers of convolution, one or more fully connected layers, and a pooling procedure make up the CNN autoencoder. The decoder is proportionately made up of convolutional and upsampling layers after either two or three fully linked layers. In contrast to the fully connected autoencoder, the CNN autoencoder functions on a series of R includes vectors rather than just one. To keep things simple, let's look at an encoder and a decoder like the ones in Fig. 2 that have one convolutional layer and one fully linked layer. The convolutional component of the encoder works with an input matrix Y ∈ R B×R, wherein R is the sequence's frames count. The layer yields a B × R × C tensor H, the components of which are computed as shown in Eq. (2).

$$\text{K(a,b,c)} = \text{g} \left(\sum_{m=0}^{h_{a-1}} \sum_{r=0}^{h_{b-1}} Y(a+m, b+n) h_c(m,n)\right) \quad (2)$$

where, Y(a, b) is the component at row a and column b of matrix Y, C is the total amount of kernels, g(•) is the non-linear activating operation, and $h_c$is a two-dimensional kernel of size $h_a \times h_b$. Keep in mind that because of zero-padding, K's initial measurements are the identical as Y's. Furthermore, although the operation carried out in the aforementioned equation is really a cross-correlation, the ML group generally refers to it as "convolution to operate.

Convolutional layers are frequently succeeded by a pooling process that lowers the input's complexity. The max-pooling procedure, that determines the highest value across a q × q windows, was utilized in (3):

$$\widetilde{K}\text{(a,b,c)} = \max \{\text{K(a',b',c')} : \text{a'} \in [\text{a . s, a . s+q-1}], \text{b'} \in [\text{b . s, b . s+q-1}] \quad (3)$$

where the step is located. It utilized q = 2 and s = 2 in the present study. A fully linked layer makes up the encoder's last layer.

The decoder is proportionately made up of a fully connected layer as the first layer, a layer that performs convolution, and an upsampling layer that replicates the input matrix's rows and columns using the identical factor 2, in this case that was utilized throughout pooling. In order to achieve two identical dimensions for the final output matrix as X, the system's final layer consists of one kernel and a convolutional layer with a linear activating function.

Applying additional convolutional and fully connected layers to the encoder and subsequently to the decoder would enhance the network's overall depth. In the trials, the real architecture was ascertained employing a validation set.

D. *Employing Genetic Algorithm for Availability Enhancement*

Towards strengthening the availability of production systems, the study deliberately employs Genetic Algorithms (GAs) as a sophisticated optimization method. Fundamentally, the main issue being addressed is the necessity of a methodical and flexible strategy to improve system availability by means of dynamic configuration optimization. Inspired by the ideas of evolution and natural selection, genetic algorithms are a powerful tool for negotiating the large solution space present in production systems' complexity. Potential system configurations are encoded into a population of people inside this optimization framework, each of which represents a distinct solution to the optimization issue. These solutions' fitness is carefully assessed using predetermined goals that are especially designed to improve system availability. This comprehensive strategy guarantees that the evolutionary algorithm converges to configurations that optimize the production system's overall availability while simultaneously reducing the likelihood of failures.

People in the population go through selection, crossover, and mutation processes as part of an iterative process known as the optimization journey. Individuals are chosen for development based on their fitness, and genetic material is transferred through crossover to produce progeny. Stochastic variations brought about by mutation encourage variety among the population. Motivated by availability-related goals, the fitness evaluation serves as a compass, pointing the algorithm in the direction of configurations that are reliable in reducing the likelihood of failure. The Genetic Algorithm becomes a powerful tool for converging across multiple generations to achieve configurations that dynamically adapt to the changing operational landscape, which eventually results in a production system optimized for increased resilience and availability.

Every individual in the population or possible solution is represented by a chromosome, frequently in binary form. Y_i, where*i* is the person's index, can be thought of as an individual solution in Eq. (4).

$$Y_i = (y_{i1}, y_{i2}, y_{i3}, \ldots, y_{in}) \quad (4)$$

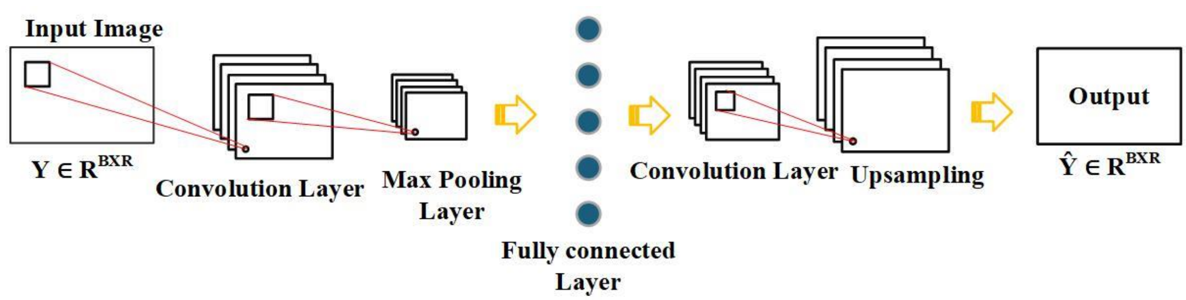

Fig. 2. Architecture of CNN Autoencoder.

The fitness of the individual is assessed by the objective function $f(Y_i)$based on the problem-specific goals. The measure of an individual's performance in relation to the optimization criteria is represented by this function.

The likelihood of choosing a person for development is directly correlated with their level of fitness. Proportional selection is a popular technique of selection in which candidates who are more fit have a larger probability of being chosen as shown in Eq. (5).

$$Q\ (Selection_i) = \frac{f(Y_i)}{\sum_{j=1}^{N} f(Y_j)} \tag{5}$$

Crossover generates offspring by fusing the genetic material of two parent solutions. Genetic material is exchanged between parents to form the offspring, and the crossover point is randomly selected as in Eq. (6).

$$Offspring_h = (y_{h1}, y_{h2} \dots, y_{hi}, \dots, y_{hn}) \tag{6}$$

A mutation modifies a person randomly; usually, this is done by flipping binary representation bits as shown in Eq. (7). The possibility of a mutation happening is determined by the mutation probability, or $Q_{mutation}$.

$$Mutation_h = (y_{h1}, y_{h2} \dots, \tilde{y}_{hi}, \dots, y_{hn}) \tag{7}$$

where, $\tilde{y}_{hi}$ is the mutated bit.

The genetic procedures result in the creation of a new population. To guarantee that the best answers are kept, the members of the new population replace the members of the old population. Until a termination requirement is satisfied, the algorithm iterates continuously. A maximum number of generations, reaching a particular fitness level, or convergence are examples of common requirements. Through these iterations, genetic algorithms gradually evolve the population in the direction of ideal solutions. The particulars and parameters, including crossover and mutation rates, vary depending on the nature of the optimization task at hand and should be adjusted accordingly. The algorithm for CNN Autoencoder-GA is given below:

Algorithm 1: CNN Autoencoder-GA

Import the required data
Describe the CNN Autoencoder architecture
Develop the CNN Autoencoder
Specify the goal function that the genetic algorithm will optimize
Analyze the system's performance using the specified configuration
Provide a fitness score based on availability and additional pertinent data
Set the Genetic Algorithm up initially
Utilize the Genetic Algorithm to maximize system settings
For increased availability, apply the optimized configuration to the production system

V. RESULTS AND DISCUSSION

Through the use of a data-driven approach, this study presents a unique way for improving production system availability and failure detection. The suggested architecture uses CNN Autoencoder for accurate failure detection and makes use of state-of-the-art techniques like deep learning and evolutionary algorithms. Overall availability is increased by using genetic algorithms to optimize system settings. The model recognizes intricate patterns after being trained on a variety of datasets, and evolutionary techniques guarantee system configuration flexibility. The model, which is implemented in Python, highlights the effectiveness of the data-driven approach in identifying and averting issues, reducing downtime, and optimizing production efficiency.

A. Model Accuracy

The degree of agreement between the deep learning model's predictions and the actual results in terms of failure detection and availability enhancement in a production system is known as model accuracy. It measures the model's efficiency in accurately detecting malfunctions and maximizing system availability. The accuracy statistic measures the percentage of cases that are properly categorized out of all the instances that the model evaluates. A high model accuracy suggests a stable and dependable performance, demonstrating the effectiveness of the data-driven, deep learning approach in conjunction with genetic algorithms in accurately predicting failure events and improving production system availability overall.

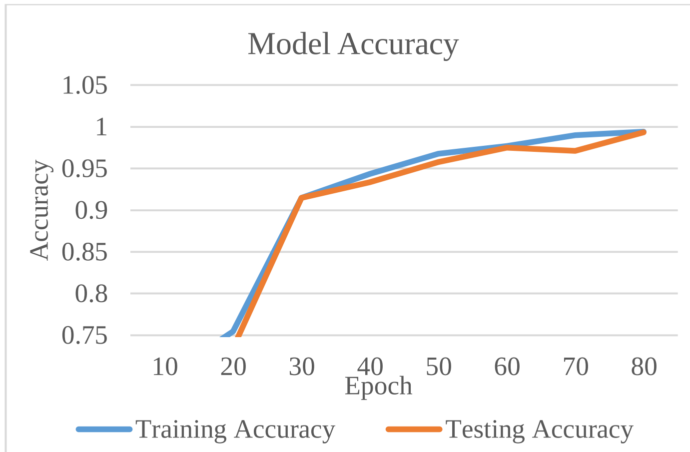

Fig. 3. Model accuracy.

The visual model accuracy graph illustrated in Fig. 3 how well the deep learning-based, data-driven strategy performed in improving production system performance. Any upward trend on the accuracy graph demonstrates an increase in the model's capacity to accurately forecast failure occurrences and raise system availability. In the context of failure detection and availability enhancement, fluctuations or plateaus may indicate regions that require extra data collection or more optimization to reach greater levels of accuracy.

B. Model Loss

When it comes to failure detection and availability enhancement in a production system, model loss is the quantitative measure of how different the actual observed values are from the projected outcomes produced by the deep learning model. This measure captures the discrepancy between the model's predictions and the actual data, indicating the degree of departure or mistake in the model's predictions. One of the main goals of training procedures is to reduce the model loss, which indicates how well the model is able to capture and reflect the underlying patterns in the data. In the context of this study, minimizing model loss is essential to guaranteeing the efficacy of the data-driven, deep learning technique and the use of genetic algorithms for production system performance optimization via improving availability and failure detection.

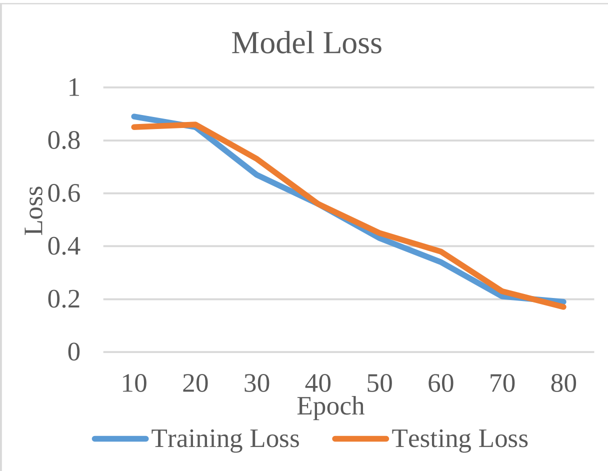

Fig. 4. Model loss.

The deep learning model's predicted mistakes are shown to have evolved in the model loss graph in Fig. 4 within the framework of improving production system performance. The model's enhanced capacity to reduce differences between expected and actual results, especially in failure detection and availability augmentation, is indicated by a declining trend in the loss graph. Finding trends or plateaus in the loss graph could inspire additional research into improving the model's architecture or training parameters to provide predictions that are more accurate.

C. ROC Curve

A graphical depiction known as the Receiver Operating Characteristic (ROC) evaluates the deep learning model's performance in terms of the trade-off between true positive rates and false positive rates.

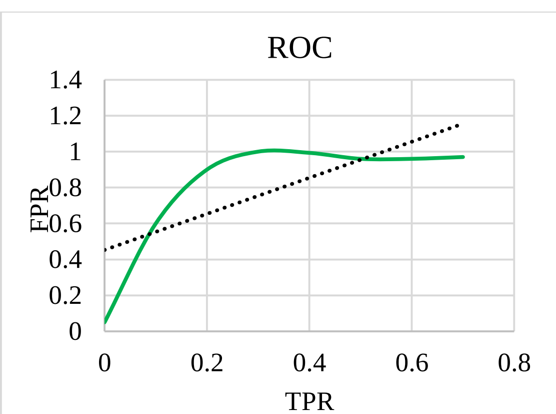

Fig. 5. ROC.

As it is depicted in Fig. 5, a thorough understanding of the model's capacity to distinguish between positive and negative occurrences linked to failure detection and availability enhancement in a production system is offered by the ROC curve, which plots the sensitivity (true positive rate) against 1-specificity (false positive rate) at various threshold settings. An increased area under the ROC curve signifies enhanced model performance in terms of accurate prediction-making while accounting for the proper ratio of true positives to false positives.

D. Performance Metrics

1) Accuracy: Accuracy is used to evaluate the system model's overall performance. Its fundamental premise is that all interactions are foreseeable. The accuracy is provided by Eq. (8).

$$Accuracy = \frac{T_{Pos} + T_{Neg}}{T_{Pos} + T_{Neg} + F_{Pos} + F_{Neg}} \tag{8}$$

2) Precision: Precision describes how comparable two or more calculations are to each other in addition to being correct. The link between accuracy and precision shows how quickly opinions may change. It is discussed in Eq. (9).

$$P = \frac{T_{Pos}}{T_{Pos} + F_{Pos}} \tag{9}$$

3) Recall: The percentage of all pertinent discoveries that were correctly sorted utilizing the procedures is known as recall. By dividing the genuine positive by the erroneously negative values, one may get the proper positive for these integers. The passage is located in Eq. (10).

$$R = \frac{T_{Pos}}{T_{Pos} + F_{Neg}} \tag{10}$$

4) F1-Score: The F1-Score computation combines recall and accuracy. To find the F1-Score, use (11), this divides the recall by the accuracy.

$$F1 - score = \frac{2 \times precision \times recall}{precision + recall} \tag{11}$$

TABLE. I MODEL PERFORMANCE

Model Performance	Percentage (%)
Accuracy	95.54
Precision	93.78
Recall	98.67
F1-Score	99.32

The model given has great efficacy in anomaly identification, as indicated by the performance indicators, which are summarized in Table I. With an accuracy of 95.54%, the model's predictions are shown to be generally accurate. With a precision of 93.78%, the model is able to correctly detect true positives among the occurrences that it classifies as anomalies. At 98.67%, the recall rate is quite high, indicating that the model is capable of accurately identifying a significant proportion of real abnormalities. Moreover, the F1-Score a balanced statistic that takes recall and accuracy into account stands out at 99.32%, highlighting the model's resilience in reaching a pleasing combination of recall and precision. Together, these measures highlight the model's excellent performance in industrial systems' real-time anomaly detection, highlighting its capacity for precise identification and reduction of false positives and false negatives.

TABLE. II COMPARISON OF PERFORMANCE METRICS

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
CNN-LSTM [26]	95.54	94.67	94.55	92.44
Bi-LSTM [27]	93.78	97.12	92.63	95.87
CNN-RNN [28]	98.67	97.35	94.69	94.26
Proposed CNN Autoencoder-GA	99.32	99.12	98.99	98.34

Table II shows the success of various techniques in the enhancing production system performance research is demonstrated by the classification performance metrics that are supplied. With an astounding accuracy of 99.32%, the Proposed CNN Autoencoder-GA notably surpasses existing models, demonstrating its capacity to accurately identify occurrences relevant to availability enhancement and failure detection.

Furthermore, a significant percentage of real positive cases are captured by the model, as seen by the high recall (98.99%) and precision (99.12%) values. The robustness of the suggested model is further shown by the F1-Score of 98.34%, which takes into account the harmonic mean of accuracy and recall. By comparison, the CNN-RNN approach has a high accuracy of 98.67%, highlighting its ability to make accurate predictions. These thorough metrics offer insightful information on the advantages of each model, with the Proposed CNN Autoencoder-GA demonstrating significant promise as a method for obtaining better results in the intended production system applications. It is depicted in Fig. 6.

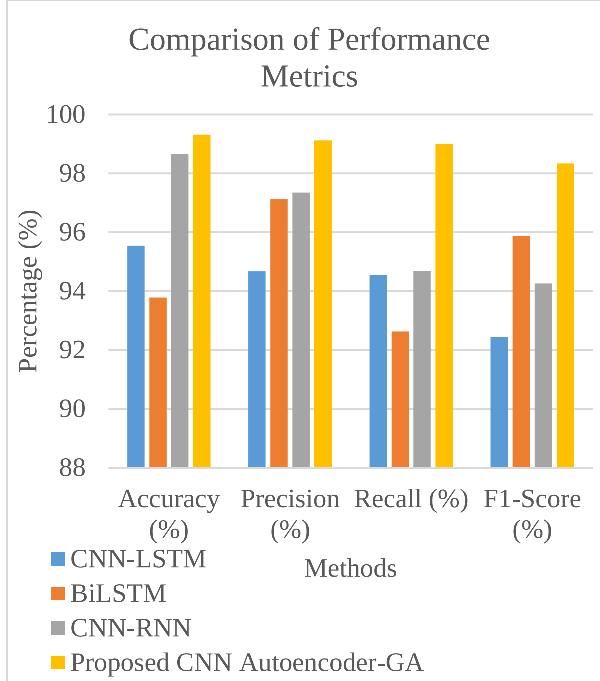

Fig. 6. Comparison of performance metrics.

E. Discussion

Through the integration of state-of-the-art methods for failure detection and availability enhancement in production systems, the suggested approach leads the way in the advancement of industrial operations. Through the use of CNN Autoencoder and deep learning, the model is able to identify complex patterns in a variety of datasets that contain data on equipment failures. A proactive approach to failure identification is made possible by this strong feature extraction capacity, which enables preventive measures to alleviate possible problems before they worsen. Furthermore, system configurations are optimized through the use of genetic algorithms, improving overall availability. With its ability to adapt to changing operating circumstances, this adaptive optimization mechanism provides a robust and effective production system management solution.

The suggested method stands out because to its exceptional accuracy of 99.32%, outperforming state-of-the-art methods such as CNN-LSTM[26], BiLSTM [27], and CNN-RNN [28]by 3.58%. This exceptional accuracy is ascribed to the complementary work of genetic algorithms and deep learning, which together offer a comprehensive approach to availability enhancement and failure detection. The suggested model, in contrast to traditional techniques, captures intricate linkages in the data, allowing for a more sophisticated comprehension of possible errors. By guaranteeing flexibility and resilience in system configurations, the application of genetic algorithms further sets the technique apart and addresses the inherent difficulties of industrial production processes. All things considered, the suggested methodology represents a major breakthrough in failure detection and availability enhancement, not only surpassing previous approaches but also providing a proactive and adaptable solution to dynamic difficulties in industrial situations.

Methodological relevance and industry prevalence served as the foundation for the benchmarking approaches chosen for this investigation. The selection of CNN-LSTM, Bi-LSTM, and CNN-RNN as typical benchmarks for well-established approaches stemmed from their extensive use in industrial settings for time-series data processing. LSTM variations such as CNN-LSTM and Bi-LSTM, which concentrate on managing temporal dependencies, tackle the sequential character of industrial data and guarantee a thorough assessment of the temporal relationship capabilities of the suggested model. Furthermore, the incorporation of CNN-based techniques takes into account the intricacy of equipment failure data, enabling a thorough evaluation of the suggested model's capacity to identify complex spatial patterns. This comprehensive comparison, which includes both CNN-based and LSTM-based approaches, offers a modern, industry-relevant framework that guarantees a careful assessment of the suggested deep learning and genetic algorithm strategy in the context of enhancing production system performance.

VI. CONCLUSION AND FUTURE SCOPE

This research has shown encouraging outcomes for improving availability and detecting failures, with a focus on using genetic algorithms. With its impressive 99.32% accuracy as well as its excellent precision, recall, and F1-Score values, the suggested CNN Autoencoder-GA model stands out as a reliable option for handling the difficulties associated with complicated production systems. The model's stability and effectiveness are further enhanced by the preprocessing stage's use of Min-Max Normalization, which guarantees that the model can absorb and analyze a variety of standardized input data. The study emphasizes how important cutting-edge methods like CNN and Autoencoder are for identifying complex patterns in real-world data, and how working in tandem with genetic algorithms improves the model's capacity to optimize for better availability and failure detection. In the future, this study will focus on extending the suggested approach's utility to other industrial contexts and investigating how well it can be tailored to real-time production scenarios. Furthermore, additional research on the interpretability of the model's decision-making procedures would improve industry acceptance and confidence in the suggested technique. The model's performance may be further enhanced by including more complex evolutionary algorithms or investigating hybrid models that incorporate other optimization methods. Furthermore, in order to support large-scale production systems, the scalability of the deep learning-based technique should be investigated. The landscape of intelligent systems for fault detection and availability enhancement in industrial settings will be significantly shaped by the ongoing improvement and adaption of these approaches as technological developments persist. All things considered, this work establishes the groundwork for a proactive and effective strategy for managing production systems by utilizing genetic algorithms in conjunction with deep learning.

The study's efficacy may be impacted by real-world unpredictability, and its generalizability to various industrial settings and datasets is restricted. The computational resources required for optimizing system setups and training the deep learning model may give rise to practical limits. The emphasis on accuracy measures obscures a thorough evaluation of the robustness of the model in different scenarios or with respect to outside influences. Furthermore, the Python software implementation might make it more difficult to integrate seamlessly with production systems that use other technology stacks.

REFERENCES

[1] A. Essien and C. Giannetti, 'A Deep Learning Model for Smart Manufacturing Using Convolutional LSTM Neural Network Autoencoders', IEEE Transactions on Industrial Informatics, vol. 16, no. 9, pp. 6069–6078, Sep. 2020, doi: 10.1109/TII.2020.2967556.

[2] W. Sun, J. Liu, and Y. Yue, 'AI-Enhanced Offloading in Edge Computing: When Machine Learning Meets Industrial IoT', IEEE Network, vol. 33, no. 5, pp. 68–74, Sep. 2019, doi: 10.1109/MNET.001.1800510.

[3] H. Wang, S. Li, L. Song, L. Cui, and P. Wang, 'An Enhanced Intelligent Diagnosis Method Based on Multi-Sensor Image Fusion via Improved Deep Learning Network', IEEE Transactions on Instrumentation and Measurement, vol. 69, no. 6, pp. 2648–2657, Jun. 2020, doi: 10.1109/TIM.2019.2928346.

[4] J. P. Yun, W. C. Shin, G. Koo, M. S. Kim, C. Lee, and S. J. Lee, 'Automated defect inspection system for metal surfaces based on deep learning and data augmentation', Journal of Manufacturing Systems, vol. 55, pp. 317–324, Apr. 2020, doi: 10.1016/j.jmsy.2020.03.009.

[5] Q. Wang, W. Jiao, and Y. Zhang, 'Deep learning-empowered digital twin for visualized weld joint growth monitoring and penetration control', Journal of Manufacturing Systems, vol. 57, pp. 429–439, Oct. 2020, doi: 10.1016/j.jmsy.2020.10.002.

[6] M. Zhang, F. Tao, and A. Y. C. Nee, 'Digital Twin Enhanced Dynamic Job-Shop Scheduling', Journal of Manufacturing Systems, vol. 58, pp. 146–156, Jan. 2021, doi: 10.1016/j.jmsy.2020.04.008.

[7] K. Alexopoulos, N. Nikolakis, and G. Chryssolouris, 'Digital twin-driven supervised machine learning for the development of artificial intelligence applications in manufacturing', International Journal of Computer Integrated Manufacturing, vol. 33, no. 5, pp. 429–439, May 2020, doi: 10.1080/0951192X.2020.1747642.

[8] 'Electronics | Free Full-Text | Artificial Intelligence-Based Decision-Making Algorithms, Internet of Things Sensing Networks, and Deep Learning-Assisted Smart Process Management in Cyber-Physical Production Systems'. Accessed: Nov. 23, 2023. [Online]. Available: https://www.mdpi.com/2079-9292/10/20/2497

[9] M. Xia, H. Shao, D. Williams, S. Lu, L. Shu, and C. W. de Silva, 'Intelligent fault diagnosis of machinery using digital twin-assisted deep transfer learning', Reliability Engineering & System Safety, vol. 215, p. 107938, Nov. 2021, doi: 10.1016/j.ress.2021.107938.

[10] R. Rai, M. K. Tiwari, D. Ivanov, and A. Dolgui, 'Machine learning in manufacturing and industry 4.0 applications', International Journal of Production Research, vol. 59, no. 16, pp. 4773–4778, Aug. 2021, doi: 10.1080/00207543.2021.1956675.

[11] Y. Li, X. Du, F. Wan, X. Wang, and H. Yu, 'Rotating machinery fault diagnosis based on convolutional neural network and infrared thermal imaging', Chinese Journal of Aeronautics, vol. 33, no. 2, pp. 427–438, Feb. 2020, doi: 10.1016/j.cja.2019.08.014.

[12] 'Sensors | Free Full-Text | Bearing Fault Diagnosis of Induction Motors Using a Genetic Algorithm and Machine Learning Classifiers'. Accessed: Nov. 23, 2023. [Online]. Available: https://www.mdpi.com/1424-8220/20/7/1884

[13] J. Gao, H. Wang, and H. Shen, 'Task Failure Prediction in Cloud Data Centers Using Deep Learning', IEEE Transactions on Services Computing, vol. 15, no. 3, pp. 1411–1422, May 2022, doi: 10.1109/TSC.2020.2993728.

[14] 'WaveletKernelNet: An Interpretable Deep Neural Network for Industrial Intelligent Diagnosis | IEEE Journals & Magazine | IEEE Xplore'. Accessed: Nov. 23, 2023. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9328876

[15] L. Sun and F. You, 'Machine Learning and Data-Driven Techniques for the Control of Smart Power Generation Systems: An Uncertainty Handling Perspective', Engineering, vol. 7, no. 9, pp. 1239–1247, Sep. 2021, doi: 10.1016/j.eng.2021.04.020.

[16] B. Farooq, J. Bao, J. Li, T. Liu, and S. Yin, 'Data-Driven Predictive Maintenance Approach for Spinning Cyber-Physical Production System', J. Shanghai Jiaotong Univ. (Sci.), vol. 25, no. 4, pp. 453–462, Aug. 2020, doi: 10.1007/s12204-020-2178-z.

[17] M.-A. Filz, J. E. B. Langner, C. Herrmann, and S. Thiede, 'Data-driven failure mode and effect analysis (FMEA) to enhance maintenance planning', Computers in Industry, vol. 129, p. 103451, Aug. 2021, doi: 10.1016/j.compind.2021.103451.

[18] H. Chowdhury, 'Semiconductor Manufacturing Process Improvement Using Data-Driven Methodologies'. Preprints, Oct. 07, 2023. doi: 10.20944/preprints202310.0056.v2.

[19] A. Kara, 'A data-driven approach based on deep neural networks for lithium-ion battery prognostics', Neural Comput & Applic, vol. 33, no. 20, pp. 13525–13538, Oct. 2021, doi: 10.1007/s00521-021-05976-x.

[20] M. Ghahramani, Y. Qiao, M. C. Zhou, A. O'Hagan, and J. Sweeney, 'AI-based modeling and data-driven evaluation for smart manufacturing processes', IEEE/CAA Journal of Automatica Sinica, vol. 7, no. 4, pp. 1026–1037, Jul. 2020, doi: 10.1109/JAS.2020.1003114.

[21] S.-K. S. Fan, W.-K. Lin, and C.-H. Jen, 'Data-driven optimization of accessory combinations for final testing processes in semiconductor manufacturing', Journal of Manufacturing Systems, vol. 63, pp. 275–287, Apr. 2022, doi: 10.1016/j.jmsy.2022.03.014.

[22] J. Friederich and S. Lazarova-Molnar, 'Towards Data-Driven Reliability Modeling for Cyber-Physical Production Systems', Procedia Computer Science, vol. 184, pp. 589–596, Jan. 2021, doi: 10.1016/j.procs.2021.03.073.

[23] J. Wu, K. Hu, Y. Cheng, H. Zhu, X. Shao, and Y. Wang, 'Data-driven remaining useful life prediction via multiple sensor signals and deep long short-term memory neural network', ISA Transactions, vol. 97, pp. 241–250, Feb. 2020, doi: 10.1016/j.isatra.2019.07.004.

[24] A. Choudhary, T. Mian, and S. Fatima, 'Convolutional neural network based bearing fault diagnosis of rotating machine using thermal images', Measurement, vol. 176, p. 109196, May 2021, doi: 10.1016/j.measurement.2021.109196.

[25] 'Machine Failure Prediction | Kaggle'. Accessed: Nov. 21, 2023. [Online]. Available: https://www.kaggle.com/code/bhaveshjain1612/machine-failure-prediction

[26] 'Application of CNN-LSTM in Gradual Changing Fault Diagnosis of Rod Pumping System'. Accessed: Nov. 23, 2023. [Online]. Available: https://www.hindawi.com/journals/mpe/2019/4203821/

[27] 'DeepRan: Attention-based BiLSTM and CRF for Ransomware Early Detection and Classification | Information Systems Frontiers'. Accessed: Nov. 23, 2023. [Online]. Available: https://link.springer.com/article/10.1007/s10796-020-10017-4

[28] 'Acoustic Anomaly Detection of Mechanical Failure: Time-Distributed CNN-RNN Deep Learning Models | SpringerLink'. Accessed: Nov. 23, 2023. [Online]. Available: https://link.springer.com/chapter/10.1007/978-981-19-3923-5_57.