Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 9, No. 5, 1040-1056
2025
Publisher: Learning Gate
DOI: 10.55214/25768484.v9i5.7072

Fault prediction models for big data analysis in industrial internet of things: A literature review

Yang Zhou[1], Nor Shahniza Binti Kamal Bashah[2*]
[1]School of Artificial Intelligence and Manufacturing, Hefei College of Finance and Economics, China; zhouyangdh@gmail.com (Y.Z.)
[2]College of Computing, Informatics and Mathematics, Universiti Teknologi MARA,40450 Shah Alam, Selangor, Malaysia; shahniza@uitm.edu.my (N.S.B.K.B.)

Abstract: Over time, fault prediction methods have seen significant advancements, becoming integral to ensuring the seamless operation of complex modern systems. These innovations are vital for controlling costs and improving safety in industries that rely on high-value assets, including sectors like oil and gas, manufacturing, and power generation. This paper primarily aims to investigate the evolution of various fault prediction techniques, highlighting their real-world applications and benefits across different industries. It provides an in-depth and up-to-date review of state-of-the-art fault detection models, focusing particularly on their deployment in industrial environments. To achieve this, advanced bibliometric approaches are used to analyze over 500 peer-reviewed articles published after 2010. A preliminary exploratory analysis identifies key players, influential authors, leading countries, and other key metrics. Additionally, a co-citation network analysis is performed to uncover and visualize key research clusters in the field. A thorough content analysis of the 100 most-cited papers follows, aiming to chart the evolution of fault detection strategies and the growing influence of artificial intelligence-based algorithms in various industrial domains. The study's findings offer valuable insights into the historical progression of fault detection methods, emphasizing their reliability while highlighting the increasing importance of intelligent algorithms. This research presents a unique perspective on the potential future directions of fault detection, providing a roadmap for researchers, policymakers, and industries with significant asset dependencies.

Keywords: *Big data analysis, Fault prediction, Industrial internet of things (IIoT), Predictive maintenance.*

1. Introduction

Fault prediction plays a pivotal role in improving asset reliability and minimizing the financial impact of equipment failures. Recently, the advent of deep learning models, particularly Long Short-Term Memory (LSTM) networks, has brought significant advancements in fault prediction, especially in scenarios involving large-scale time-series data. LSTM networks are highly effective in capturing long-term dependencies in data, making them an ideal choice for predicting future failures based on historical data patterns. LSTM's ability to handle sequential and non-linear relationships is essential for real-time fault detection in industries like oil and gas, manufacturing, and power generation, where operational efficiency and safety are paramount. As industries that heavily rely on assets, such as oil and gas, manufacturing, and power generation, continue to grow, the demand for systems that enhance productivity, availability, and efficiency with minimal maintenance costs is escalating. In the early stages of industrial development, fault detection posed significant challenges. However, the integration of computational intelligence and machine learning (ML) has facilitated the creation of advanced solutions for predicting machinery faults with greater accuracy. Although the terms "fault" and "failure" are often used interchangeably in industrial contexts, they describe different conditions of equipment. A

History: Received: 24 February 2025; Revised: 16 April 2025; Accepted: 21 April 2025; Published: 12 May 2025
* Correspondence: shahniza@uitm.edu.my

fault refers to a situation where equipment is still operational but exhibits reduced functionality, or fails to perform as expected when needed. In contrast, a failure occurs when the equipment completely ceases to function. For example, when a valve fails to open or close as commanded, it is considered a failure, while a valve that leaks but continues to function is categorized as a fault. Faults typically allow for a period of time in which maintenance can be planned, helping ensure continuous operation, safety, reliability, and optimal maintenance costs.

Early fault detection is crucial to prevent failures, control repair costs, and improve asset reliability. Identifying faults early allows for timely corrective actions, which help avoid costly breakdowns—repair costs for unplanned failures are often several times higher than for planned maintenance. Higher reliability, reflected by the mean time between failures, signifies a reduced probability of failure over time, improving the equipment's ability to perform its intended function. Additionally, early fault detection aids in better maintainability by providing more time for maintenance planning, reducing downtime, and ensuring faster repairs. This ultimately enhances both plant availability and overall system performance. Reliability, maintainability, and availability are interrelated, with improvements in reliability and reduced repair time leading to increased availability.

Several literature reviews have addressed fault prediction models, often focusing on specific applications, such as wind turbines, power inverters, and software quality, among others. While earlier reviews have provided insights into fault prediction methodologies, most have primarily examined statistical approaches and did not employ a systematic literature review framework. Recent studies, such as those by Alauddin, et al. [1] have utilized bibliometric analysis (BA) tools to conduct an SLR, but this paper distinguishes itself by focusing on AI-based models for fault prediction, with particular attention to research from the past decade. Traditional statistical methods, such as ARIMA models and hidden Markov models (HMM), were once widely used tools. The ARIMA model is used in time series analysis to capture the trend of equipment failures and predict future failure modes. The HMM model can simulate the changes in system states, provide transition probabilities between different health states, and thus predict failures. However, with the rise of the Industrial Internet of Things (IIoT) and big data analysis, methods based on AI and machine learning have gradually become mainstream to improve prediction accuracy. In order to meet the challenges of large-scale data processing, distributed computing technologies, such as Apache Spark, are being widely used to process large amounts of data from IIoT devices to achieve real-time data processing and fault prediction.

The SLR presented in this paper leverages advanced bibliometric analysis tools to perform co-citation analysis and identify clusters of research within the field. By applying multivariate clustering methods, this study visualizes these research clusters and identifies key themes and trends. A detailed content analysis of the top 100 most-cited papers is provided, focusing on AI techniques and their applications in industrial settings. In contrast to previous reviews, this paper offers a comprehensive, up-to-date SLR, utilizing modern bibliometric techniques to analyze the literature on fault prediction. The SLR framework used in this study avoids the biases often associated with traditional reviews, offering a more objective and holistic view of the research landscape. Key findings from this analysis will contribute to identifying future research directions and establishing a roadmap for further advancements in the field.

The structure of this paper is organized as follows: Section 2 outlines the conceptual framework for the systematic literature review (SLR) and details the criteria used to select relevant studies from key academic databases. Section 3 provides an exploratory bibliometric analysis to uncover key research indicators, such as the most cited papers, leading authors, and top-performing countries. In Section 4, a co-citation network analysis is conducted to highlight the relationships between significant studies. Section 5 presents an in-depth content analysis of the 100 most-cited articles. Section 6 examines current trends in fault detection research, while Section 7 offers a summary of key findings and suggests potential avenues for future research.

2. Conceptual Framework and Literature Search Strategy

In the process of modeling fault prediction, classical statistical methods such as ARIMA and HMM have been traditionally used for time-series analysis and system state prediction. However, these models face challenges when handling non-linear patterns and large-scale datasets. Recently, Long Short-Term Memory (LSTM) networks, a type of recurrent neural network (RNN), have gained traction in fault prediction models. LSTM networks are particularly effective in capturing long-term dependencies within time-series data, making them suitable for fault prediction in industrial systems with complex, non-linear behavior. By learning the temporal relationships in equipment performance, LSTM networks can predict faults more accurately and dynamically adapt to changing operational conditions. Especially in the processing of large-scale data sets, Long Short-Term Memory (LSTM) are used to mine research topics in large-scale literature, thereby revealing the evolution trend in the field of fault prediction. LSTM can help researchers identify the most important research directions and technological advances in fault detection by performing topic modeling on a large amount of literature data. In addition, hidden Markov models (HMMs) are also widely used in some fault prediction systems, especially in those cases where the system state changes. HMMs can effectively simulate the transition process between different health states, thereby providing accurate fault prediction. Traditional time series models, such as ARIMA, have advantages in capturing the long-term trend of equipment failures, but often face the limitation of not being able to handle nonlinear data and complex patterns. With the application of machine learning methods in fault prediction, hyperparameter optimization has become a key step to improve the accuracy of model prediction. Common hyperparameter optimization methods, such as grid search, random search, and Bayesian optimization, can find the best configuration among different model parameters, thereby significantly improving the accuracy and stability of fault prediction. By tuning the hyperparameters of machine learning models such as neural networks and support vector machines, more accurate fault prediction results can be obtained. On the other hand, with the rise of big data, distributed computing technologies such as Apache Spark provide powerful data processing capabilities for real-time fault prediction. Using Apache Spark, researchers can efficiently process and analyze data from multiple sensors in a distributed environment, thereby achieving real-time updates and dynamic optimization of fault prediction models (Figure 1).

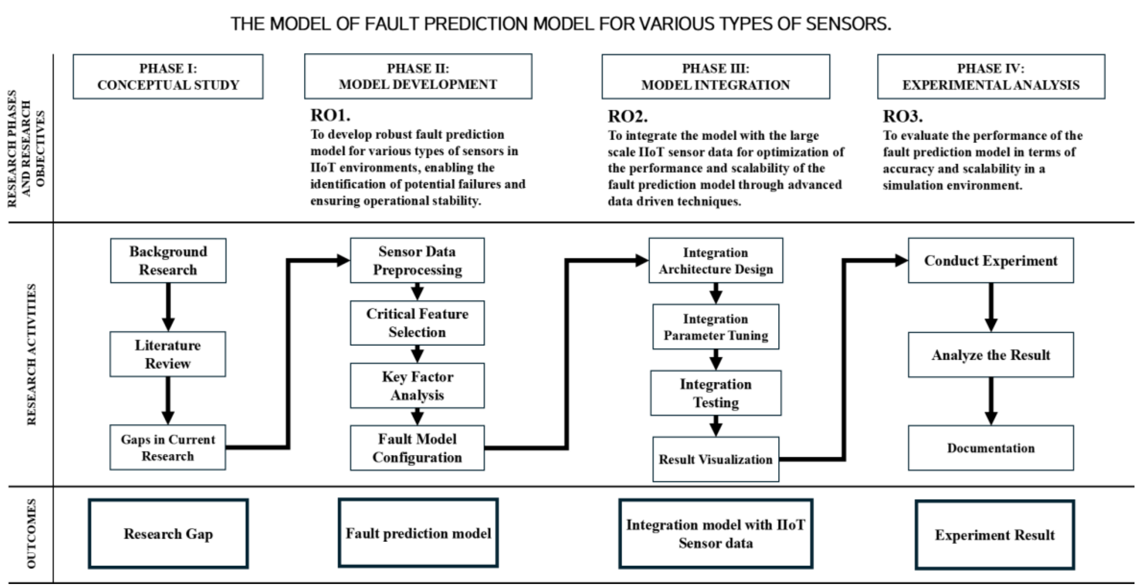

Figure 1.
Research Design Framework for Fault Prediction Model.

Systematic literature reviews (SLRs) are typically carried out following a structured process that begins with identifying the most relevant research through targeted keyword searches. In this study, we conducted comprehensive searches across two prominent research databases: Web of Science (WOS) and SCOPUS. SCOPUS is known for being the largest repository of peer-reviewed academic articles, while WOS holds the distinction of offering the longest history of peer-reviewed research, dating back to 1900. The search queries utilized included combinations of keywords such as “Fault” and “Reliability,” as well as “Machine Learning” and “Artificial Intelligence,” ensuring that the search captured all relevant research related to fault detection, reliability, and advanced analytical methods like ML and AI.

To refine the search, we focused on research from the past decade, specifically from 2010 to 2020, and limited the selection to peer-reviewed journal articles and conference papers published in English. After gathering the relevant documents, we meticulously screened them to eliminate duplicates and irrelevant studies. This resulted in a dataset of 552 documents from the SCOPUS database were peer-reviewed journal articles. These selected papers were then analyzed by advanced bibliometric analysis (BA) tools to identify trends and key contributions in the field.

3. Exploratory Bibliometric Analysis

In the co-citation network analysis, we found that the early research on HMM and ARIMA models still occupies an important position in the field of fault prediction, particularly for time series data analysis. However, with the rise of deep learning, Long Short-Term Memory (LSTM) networks have increasingly replaced traditional statistical models for time-series forecasting and fault detection. LSTM's unique ability to capture long-term dependencies and handle complex, non-linear data patterns has made it a preferred choice in fault prediction, especially in real-time industrial applications. This shift in research focus towards deep learning, and particularly LSTM, reflects the growing demand for more accurate and adaptive models in industries where large-scale data and real-time analysis are critical. These classic statistical methods provided the basic framework for fault prediction models in the early days, but with the advancement of machine learning technology, more and more research has turned to using techniques such as deep learning and support vector machines (SVM) to deal with complex nonlinear problems. HMM has a unique advantage in simulating changes in the health state of the system, while ARIMA predicts future failure trends by analyzing the historical failure data of the equipment. Although these methods are still effective, they are generally unable to handle large-scale data and complex multi-dimensional features.

A systematic literature review (SLR) serves as an organized method for exploring research in a particular field or topic. It helps to establish an understanding of the development and benefits of the research over time, offering insights into how methodologies and models evolve. This process not only provides a comprehensive overview of the field but also helps shape the direction for future research. An SLR involves several stages, one of which is exploratory analysis, aimed at understanding key characteristics such as the growth of publications, influential works, leading countries in the field, and the most commonly used keywords. In this study, bibliometric analysis was carried out using the “Bibliometrix” tool, which is designed for descriptive bibliometric analysis (Figure 2).

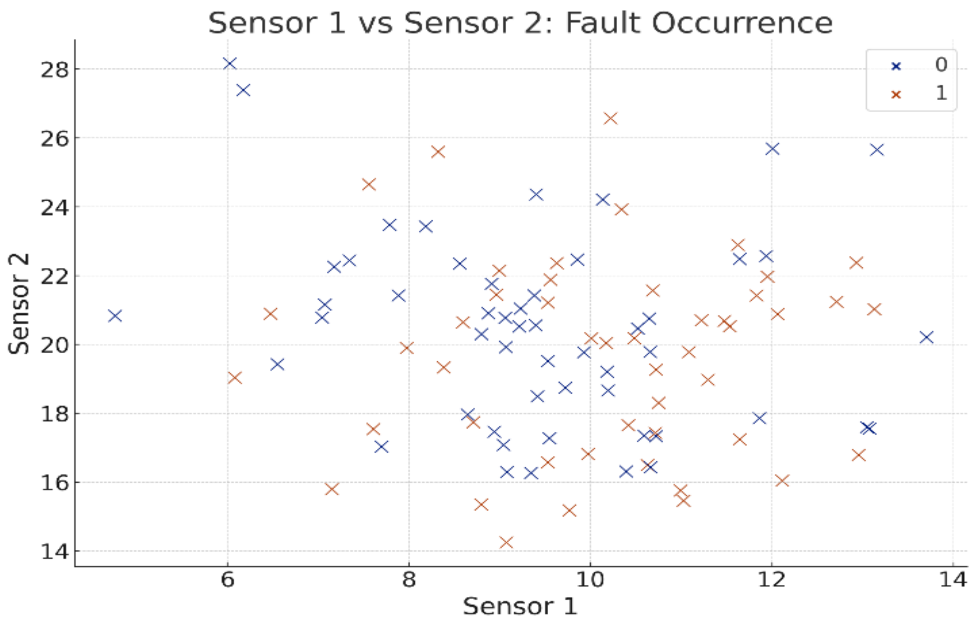

Figure 2.
Sensor 1vs Sensor 2: Fault Occurrence.

3.1. Growth of Publications

The growth of publications over the last decade is illustrated. The analysis spans the period from 2010 to 2020, with the most recent complete year, 2019, showing approximately 141 published articles. In the early part of 2020, there were around 10 articles published in just the first week. To model the publication trend, both linear and second-order polynomial regression models were applied. The linear regression model indicated an annual growth rate of roughly 18% from 2010 to 2014, which accelerated to about 42% between 2015 and 2019. The second-order polynomial regression model demonstrated a superior fit, achieving a coefficient of determination (R^2) of 0.973, in contrast to the linear model's R^2 of 0.845. The analysis highlighted a consistent annual growth rate of approximately 2.57%, suggesting a rising interest in artificial intelligence (AI) and machine learning (ML) techniques for fault detection. This growth in publications can be attributed to the enhanced performance of AI- and ML-based approaches in solving practical issues, as well as significant improvements in technology and computational capabilities in recent years.

Although the roots of fault detection using AI and ML trace back several decades, the past 10 years have witnessed a marked increase in research activity in this area. This growth is primarily attributed to factors such as technological advancements, the rise of data analytics, greater computing capabilities, and the growing emphasis on maintenance cost optimization in industries. Notably, an exponential rise in the number of publications in the last five years suggests that these AI- and ML-driven approaches have become a central focus in fault detection research.

3.2. Influential Authors and Publications

This section delves into the most impactful authors and publications in the field of fault prediction models for the Industrial Internet of Things (IIoT) from 2010 to 2020. The authors and their corresponding publications are highlighted in the table. Global citations (GC) refer to the citation count of a paper across all databases, as retrieved from SCOPUS at the time of data collection. On the other hand, local citations (LC) indicate how often a publication is referenced within the specific collection of articles analyzed (Figure 3).

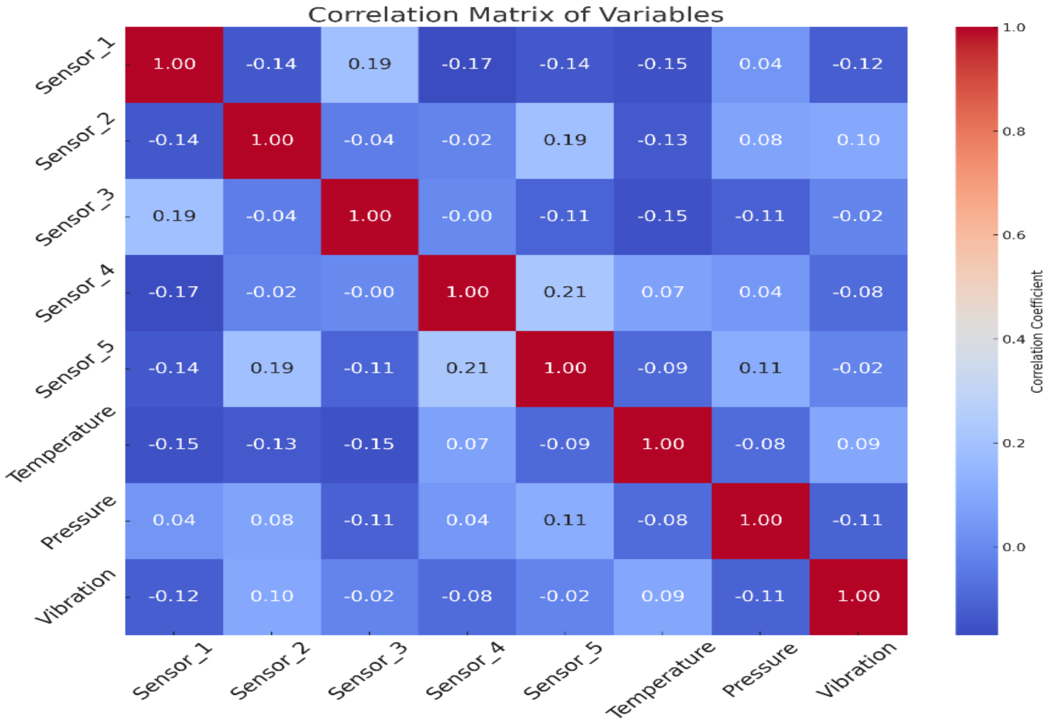

Figure 3.
Correlation Matrix of Variables.

Among the various scholarly works, the 2013 paper authored by Tamilselvan and Wang [2] has accumulated the most substantial number of worldwide references, totaling 273. This investigation delves into the utilization of Deep Belief Networks for classifying health states, aiming to boost system dependability and curtail maintenance expenditures. In 2015, a study by Li, et al. [3] cited 142 times, proposed a multimodal deep support vector machine method for the diagnosis of gearbox anomalies. Their findings underscored the efficacy of deep learning as a potent tool for refining gearbox fault identification and enhancing the overall robustness of the system. In Bacha, et al. [4] and associates employed support vector machines (SVM) for diagnosing faults in power transformers, specifically targeting the analysis of dissolved gases. Their results illustrated that the proposed methodology significantly elevated the accuracy of fault detection and increased the operational efficiency of transformers.

Additionally, It reveals the authors with the highest number of publications. Li, et al. [5] lead with 16, 12, and 11 publications, respectively.

3.3. Leading Countries in Research Productivity

The application of AI and ML in fault diagnosis through intelligent algorithms has gained substantial attention due to its potential to address challenges faced by real-world industrial systems. While this field has its roots dating back to the 1970s, the last decade has seen considerable advancements driven by technological progress, growth in data analytics, increased computing power, and the rising importance of maintenance cost management for industries. Notably, the past five years have witnessed a sharp rise in research activity within this domain (Figure 4).

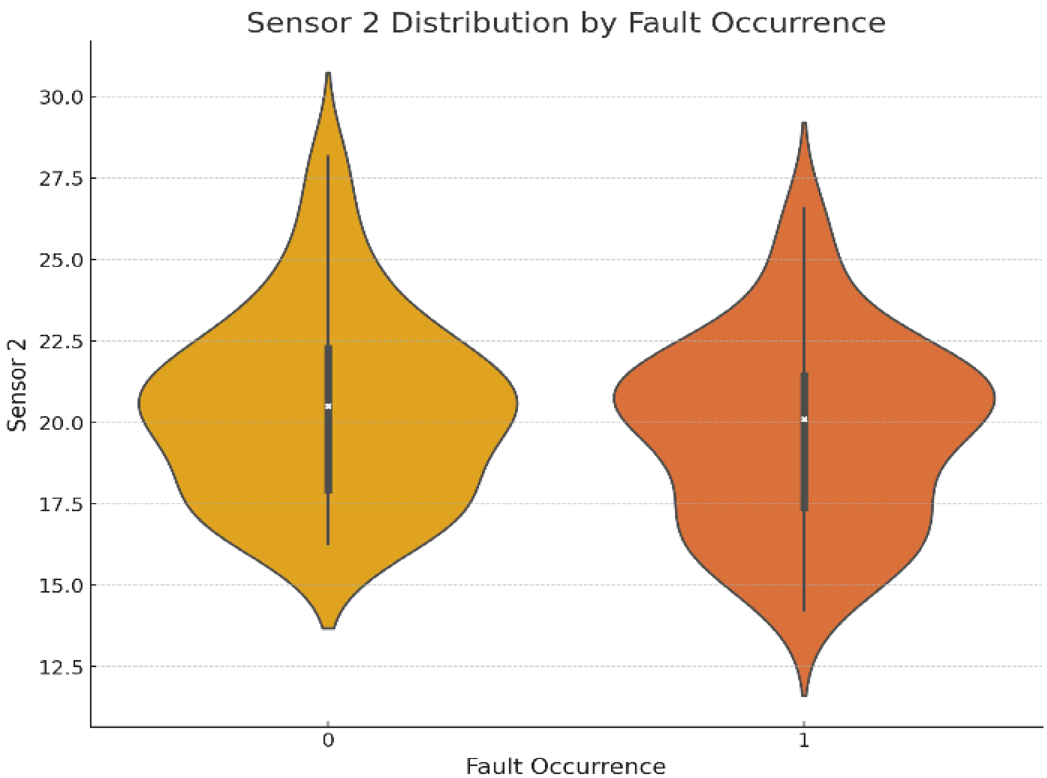

Figure 4.
Sensor 2 Distribution by Fault Occurrence.

It illustrates the research output by country, showing the frequency of publications. The countries with the highest number of publications are China, the USA, and India respectively. China's output is approximately 4.6 times greater than that of the USA, while the gap between the USA and India is around 1.9 times. To assess the impact of these publications, we calculated the ratio of total citations (TC) to the number of publications (NP). Canada stands out with a ratio of 4.2, placing it at the top, followed by Turkey and the USA. Despite China's dominant publication volume, its TC/NP ratio ranks eighth globally, which suggests that the impact of its research is relatively lower in comparison to other leading nations. It also presents a global map depicting the scientific productivity in this area.

4. Network Analysis

Long Short-Term Memory (LSTM) networks are now at the forefront of predictive maintenance frameworks due to their unique ability to process sequential and time-series data effectively. Unlike traditional machine learning models like Support Vector Machines (SVM) and Artificial Neural Networks (ANN), which are often limited in capturing long-term dependencies and non-linear patterns, LSTM networks excel in learning from large volumes of temporal data. They are particularly well-suited for fault prediction in complex industrial systems, where the relationship between past system states and future failures is crucial. LSTM models can learn from historical data, adapt to new patterns, and predict future faults with greater accuracy and reliability, providing an edge over traditional methods in real-time fault detection applications. This section provides an analysis of the research network, which helps to reveal the structure of the field, the global collaborations among authors, and the recurring themes that appear across various studies. The analysis utilizes network theory, a discipline that originated in computer science, to identify patterns and influences within research domains. Specifically, a co-citation network analysis is applied to investigate how knowledge is structured within the Fault prediction domain, the main focus of this review. This approach visualizes the relationships between papers by highlighting shared topics or references.

4.1. Centrality-Based Clustering

Co-citation analysis is a powerful tool for identifying similarities among research papers. By examining the connections between the cited documents in a co-citation network, it is possible to uncover structural patterns. This method allows for the identification of key research papers and their relationships based on how frequently they are cited together. The co-citation network can be developed using bibliometric software tools, such as the "bibliometrix" R-package, which we applied for our analysis (Figure 5).

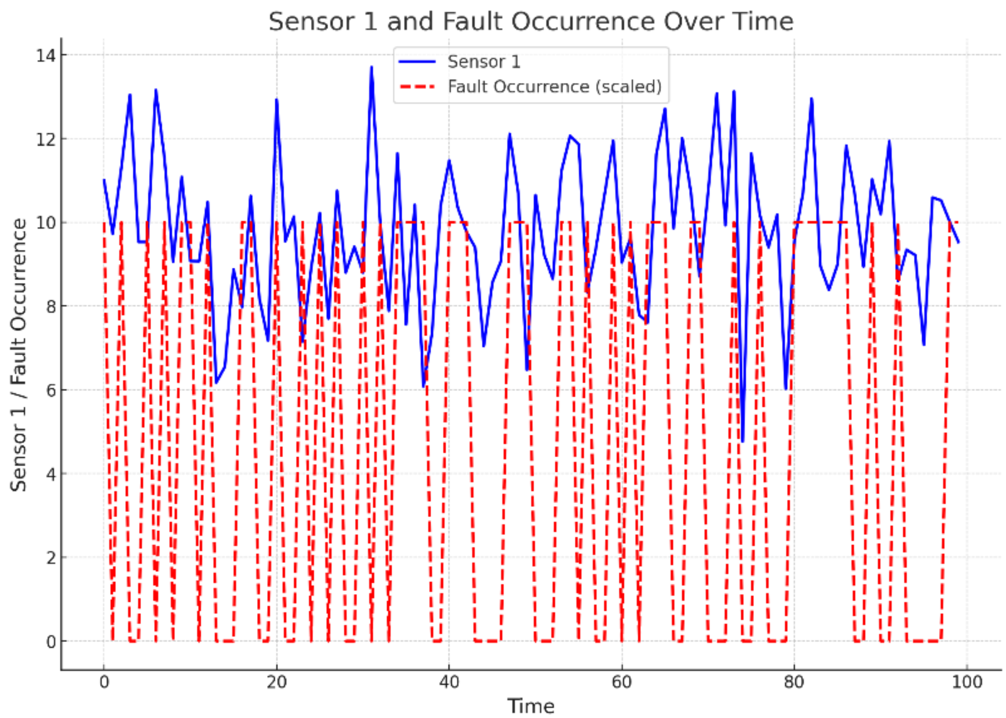

Figure 5.
Sensor 1 and Fault Occurrence Over Time.

The investigation underscores the most frequently referenced works within the network (27 sources), which are organized into four distinct groups, each represented by a unique color. The pivotal papers within each group are depicted with larger nodes, indicating their significant impact in that specific cluster. The study categorizes the four groups and provides an overview of the themes explored by these studies. Cluster 1 is centered around autonomous fault detection methods, examining techniques such as seismic analysis, neural networks, and an array of machine learning approaches. Cluster 2 is predominantly concerned with data-centric models and AI methodologies, including deep learning and support vector machines (SVMs). Clusters 3 and 4 concentrate on fault detection applications and the extraction of knowledge through AI methods, such as artificial neural networks and particle swarm optimization.

4.2. Modularity-Based Clustering

Modularity is a concept used to assess the community structure within a network, determining how well links are distributed within clusters compared to those between clusters. To analyze the

ISSN: 2576-8484
Vol. 9, No. 5: 1040-1056, 2025
DOI: 10.55214/25768484.v9i5.7072

modularity of the co-citation network, we utilized the open-source tool Gephi, which is widely used to represent and analyze complex networks. Gephi uses modularity classes to cluster the network into distinct groups. In this analysis, three modularity clusters (labeled 0, 1, and 2) were identified. Within each cluster, the top 10 references were selected based on their PageRank, a measure of their importance within the network.

It also highlights the key contributions of these influential papers. The five most highly ranked papers in each cluster were identified and summarized. In cluster 0, the top papers include works by Nimmo [6]; Gertler [7]; Massoumnia [8] and Basseville and Nikiforov [9]. Cluster 1's most prominent papers include Silver, et al. [10]; Silver, et al. [11]; Wang and Wang [12] and Zhu, et al. [13]. These papers are discussed in more detail in the following sections, where we analyze their contributions within each modularity class (Figure 6).

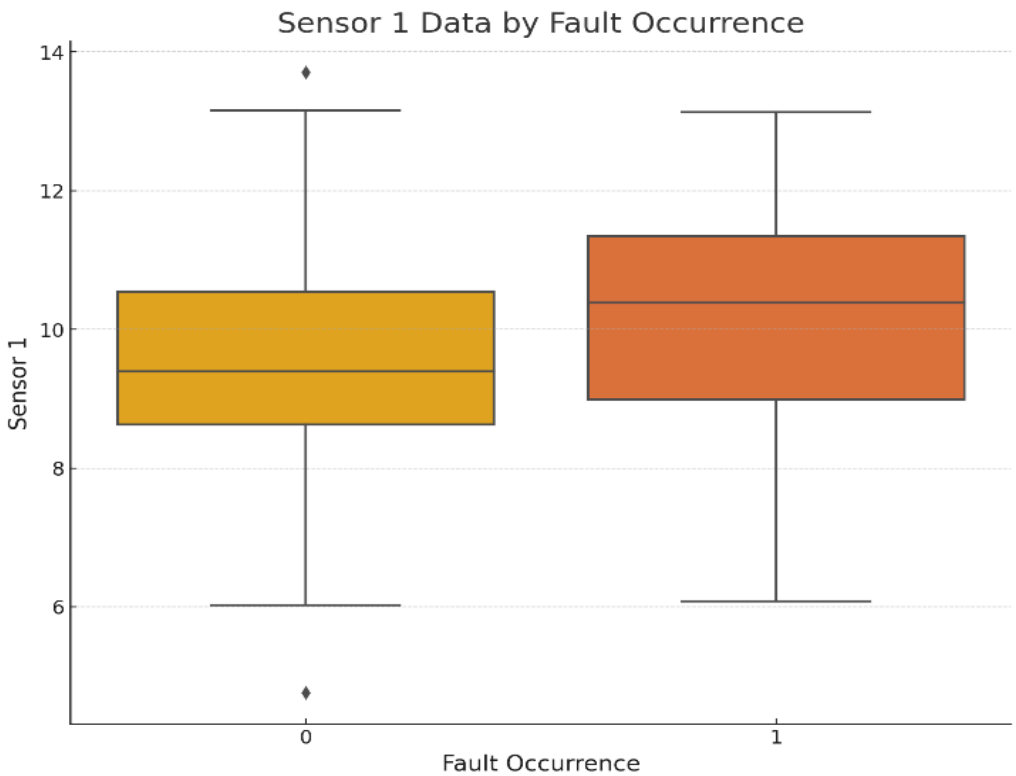

Figure 6.
Sensor 1 Data by Fault Occurrence.

4.2.1. Cluster 0: Fault Identification in Processes and Systems

This group accounts for approximately 13% of the overall papers considered in the review. In this section, we provide a brief overview of the ten most significant studies identified through PageRank, all of which concentrate on fault identification utilizing AI-driven methodologies. A prominent study from the early 1990s by Isermann [14] examines foundational fault detection methods, including parameter estimation techniques for systems operating in continuous time. Gertler [7] proposed a fault detection and isolation approach grounded in parity relations, designed to handle both additive and multiplicative faults. His research also delves into various isolation methodologies, emphasizing the importance of separating from external interferences. Chen and Patton [15] developed a fault-tolerant architecture for flight control, employing linear matrix inequalities to model fault impacts through linear-fractional transformations. Their simulation results indicated the system's stability and effective performance under fault conditions. Persis and Isidori [16] took a different approach by employing a geometric method for fault detection and isolation in nonlinear systems. In another study, Isermann [14] later highlighted the growing significance of model-based fault prediction methods, which are essential in industrial systems to enhance safety, reliability, and performance. These methods use dynamic process models to analyze input-output signals, identifying discrepancies between normal and faulty conditions. As these techniques are increasingly adopted in manufacturing and industrial environments, they offer deeper insights into faults, surpassing classical trend-based methods.

4.2.2. Cluster 1: Fault Detection Methods and Applications

This cluster makes up the largest share of papers in the co-citation network, representing more than three-quarters of the total publications. Below, we summarize the ten most impactful papers identified through PageRank, which focus on fault detection methods and their applications. One significant study by Silver and his team utilized a neural network-based supervised learning technique, enabling a system to master the game of Go through reinforcement learning, without relying on any prior human knowledge. Subsequently, Silver's group developed an AI-powered prediction system using a fully abstract model, which offered enhanced prediction accuracy compared to conventional deep neural network designs. Wang and Hu [17] proposed a nonlinear flexo-electricity model for energy harvesting, which effectively converts ambient vibrations into electrical power for microelectromechanical systems, eliminating the need for batteries. This model accounts for geometric nonlinearity and damping effects, improving the accuracy of energy harvester behavior predictions. Zhu, et al. [13] introduced a Bayesian classifier-based fault diagnosis method for transformers, integrating dissolved gas-in-oil analysis with traditional electrical testing. Their results showed that this combined approach delivered more reliable fault classification in transformers than using either method individually. Zou proposed a fuzzy logic-based fault diagnosis technique for power transformers, utilizing an innovative coding membership function. His approach showed enhanced robustness and accuracy when applied to historical data and test results. Huang and colleagues focused on extreme learning machines for feedforward neural networks, which provided fast learning speeds and superior generalization performance. Their experimental results indicated that the new algorithm outperformed traditional learning algorithms in terms of both speed and accuracy. Lastly, Wu, et al. [18] utilized support vector machines (SVM) for detecting discharge and thermal faults in transformers. Their experimental results indicated that the SVM model significantly improved fault detection rates.

4.2.3. Cluster 2: Fault Diagnosis through AI Methodologies

This cluster constitutes roughly 14% of the total papers within the co-citation network. In this segment, we offer a synthesis of the ten most impactful studies in this domain. Guardado, et al. [19] conducted a comparative analysis on the use of neural networks for fault detection in transformers. The network was trained utilizing five widely recognized criteria for identifying faults, grounded in dissolved gas analysis techniques. When tested with new gas analysis results, the network demonstrated a high success rate in diagnosing faults. Wang and Hu [17] introduced a vibration-based fault detection method using fuzzy logic for pumps. Their model employed fuzzy logic to capture the uncertainty between various fault symptoms and conditions, effectively classifying frequency spectra that represent different pump faults. The results highlighted the model's strong capability in identifying and classifying mechanical faults in machinery. Chen and Patton [15] proposed a fault classification model that combined manifold learning with an adaptive wavelet-based neural network. The model was validated using data from roller bearings and showed promising performance in fault classification. Castro and Miranda [20] along with Naresh [21] also developed neural network-based models to identify faults in transformers, further advancing the use of AI in transformer fault detection. Fei and Zhang [22] introduced a fault diagnosis technique for power transformers using a support vector machine (SVM) enhanced by a genetic algorithm. Tang and colleagues proposed a probabilistic classifier for transformer fault detection, utilizing a particle swarm optimizer to improve the accuracy of dissolved gas analysis. Additionally, Shintemirov, et al. [23] developed a genetic programming approach for feature extraction in dissolved gas analysis, using bootstrap methods to enhance fault detection in power transformers. These studies emphasize the growing use of novel AI techniques, including genetic algorithms and SVMs, for improving fault detection in power systems and transformers, yielding encouraging results.

5. Result Analysis

Looking forward, the integration of LSTM networks with the growing volume of industrial IoT data presents significant opportunities for improving fault prediction accuracy. With the continuous evolution of big data technologies, LSTM models can be enhanced by integrating them with hyperparameter optimization techniques, such as grid search and Bayesian optimization, to further fine-tune the model's performance. Additionally, transfer learning, which allows LSTM networks to leverage pre-trained models and adapt them to new industrial environments, holds great promise for improving fault detection accuracy in previously unseen operational conditions. As industries continue to deploy IoT devices at scale, LSTM networks will play an increasingly vital role in ensuring the reliability and efficiency of critical assets. This section provides a content examination through an exploratory methodology, with a temporal focus on research driven by practical applications. The primary emphasis is on the contribution of artificial intelligence technologies in sectors with significant asset dependencies, including oil, gas, and petrochemical industries. The objective is to assess the evolution of fault identification and diagnostic studies through the lens of AI and related technologies. A total of 100 highly-cited papers, published between 1989 and 2019, were reviewed to assess the scope and progression of fault detection efforts aimed at improving reliability over the last three decades. This analysis also examines the AI techniques and other methods employed in fault prediction, while exploring industry-specific applications that have led to advancements in equipment reliability and operational efficiency.

Approximately 79% of the research papers concentrate on fault prediction, highlighting the essential role of fault prediction in ensuring equipment reliability. The remaining 21% address topics such as equipment performance, fault prediction, and system and equipment classification. The first research period spans from 1989 to 1998. During this phase, the primary focus was on fault prediction, with neural networks being the most widely used AI algorithm. The studies covered diverse industrial applications, including power systems, information technology, and process systems. Period II, covering 1999 to 2008. This period saw the emergence of more sophisticated AI techniques and a broader array of industrial applications. The increasing volume of research indicates a rising demand for fault prediction solutions. The most recent period, from 2009 to the present. In this context, progress in AI, including deep learning, adaptive neural networks, and reinforcement learning, has started to demonstrate more significant potential. The scope of applications has broadened, extending beyond traditional power systems to emerging domains like information technology and renewable energy, covering a more diverse array of machinery (Figure 7).

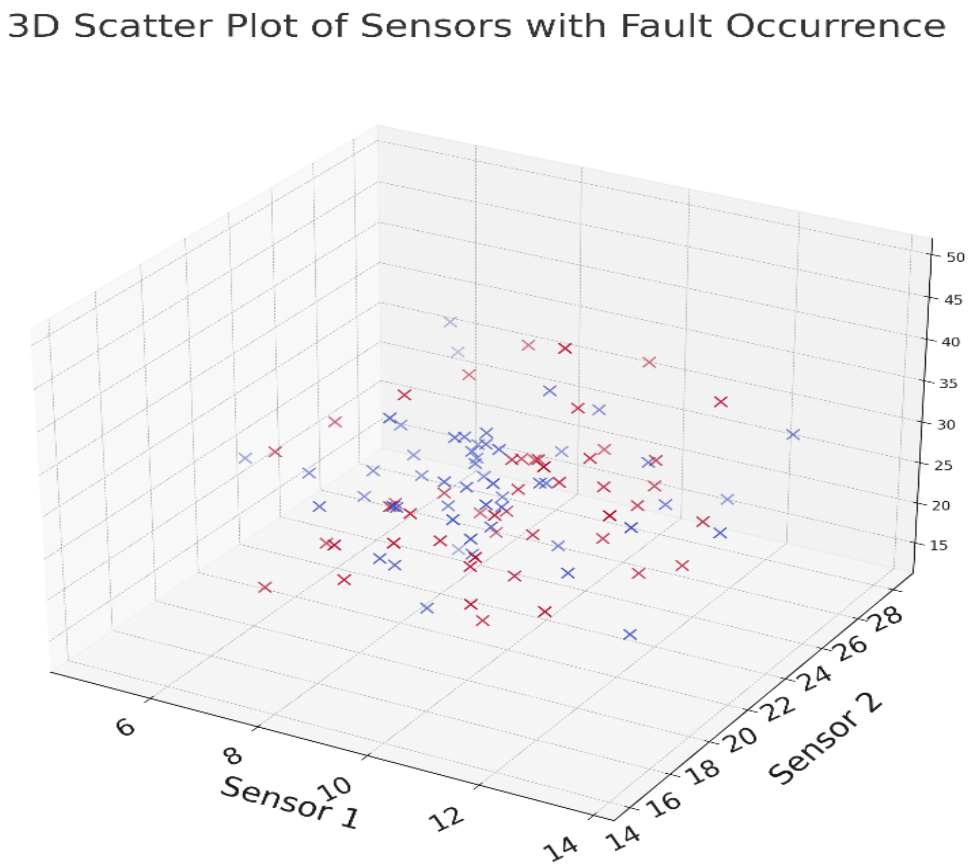

Figure 7.
3D Scatter Plot of Sensors with Fault Occurrence.

5.1. Modeling Framework

We provide an overview of AI-driven modeling frameworks and techniques commonly utilized. The distribution of widely-recognized predictive modeling methods from AI and machine learning (ML), which includes top-cited papers analyzed in this review. Neural networks (NN) and support vector machines (SVM) are the predominant methods, appearing in over three-quarters of the publications.

During the first period (from the late 1980s to the late 1990s), neural networks (NN) or artificial neural networks (ANN) were among the most frequently applied machine learning algorithms for fault prediction. One significant reason for their popularity is that ANNs possess strong self-learning capabilities, allowing them to model complex relationships within data. Fault detection and diagnostics often involve intricate issues, which NN can manage effectively while being relatively easy to set up and implement. Additionally, these algorithms automatically perform feature extraction by assigning minimal weight to irrelevant data, thus reducing the need for manual feature extraction. Approximately 70% of the studies during this period employed NNs to detect faults and address related complex challenges. However, about 30% of the research relied on alternative methods, including abstraction hierarchy models, evolutionary programming-based fuzzy systems, multiway principal component analysis (PCA), and partial least squares for fault detection, process monitoring, and high-risk component prediction in complex systems.

For example, one notable study proposed a neural network-based methodology for fault diagnosis in an oil refinery, specifically in the fluidized catalytic cracking process. This NN system was able to generalize its knowledge and detect previously unseen fault combinations without additional training. It was also capable of handling incomplete and uncertain data effectively. Another study introduced an adaptive pattern recognition approach using ANN, leveraging its high adaptation capabilities to model complex relationships between input features and fault detection times. In addition, methods like multiway PCA and partial least squares were utilized to enhance fault diagnosis in batch processes.

The second period (1999–2008) witnessed continued dominance of neural networks, often in combination with other algorithms. During this time, new techniques emerged to improve fault detection. Around half of the top-cited papers from this period still used neural networks to address fault detection issues. A comprehensive review examined the most recent AI-based diagnostic advancements for electrical machines and drives. These methods were found to be effective individually or when integrated with traditional techniques.

In addition, numerous investigative endeavors have proposed hybrid strategies, incorporating the amalgamation of Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs) for identifying faults within radial distribution grids. This method typically involved the utilization of principal component analysis (PCA) to preprocess the data, followed by fault categorization based on impedance through a synthesis of support vector classifiers and multilayer perceptron networks. In another instance, ANNs were employed to forecast and model aquatic quality parameters, with the developed model being evaluated on fresh, previously unobserved data to determine its precision. This highlighted the potential of ANNs in modeling complex ecological variables. Beyond neural networks, other intelligent algorithms were also explored for fault diagnosis. One example is a Takagi-Sugeno fuzzy model-based approach for networked control systems, which outperformed older techniques by not requiring precise knowledge of network-induced delays, and handling data packet losses more effectively. Moreover, SVM-based models demonstrated superior performance when fault samples were limited, as shown in their application to a turbo-pumping system, which highlighted SVM's ability to generalize well even with few fault samples.

In the third period (2009–2018), a significant shift was observed in research approaches, with more than 60% of the studies introducing new techniques and algorithms, while the remaining 40% continued to utilize neural networks either alone or in combination with other methods. Notably, around 65% of the papers from this period were focused on fault prediction using modern tools and advanced algorithms, highlighting the growing trend of applying novel methods. This transition emphasizes the advantages of newer algorithms compared to those used in the earlier periods. Moreover, various research efforts have suggested hybrid methods, including the fusion of Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs) for detecting faults in radial distribution networks. The approach typically involved applying principal component analysis (PCA) to process the data, followed by fault classification based on reactance. In another case, ANNs were used to predict and model water quality indicators, with the trained model being tested on new, unseen data to assess its accuracy. This demonstrated the capability of ANNs in simulating intricate environmental factors.

For instance, deep neural networks (DNNs) were employed to capture fault patterns and identify problems in rotating machinery. This method overcame several drawbacks of earlier techniques, as deep learning's intricate architectures enable the extraction of valuable insights from unprocessed data and the modeling of complex nonlinear relationships. One methodology suggested a dual-phase learning framework for diagnosing mechanical systems, where sparse filtration, a self-directed neural network with two strata, extracted characteristics from vibrational signals, and soft-max regression was employed to categorize the machinery's operational health based on these extracted traits. In an alternative method, convolutional neural networks were harnessed for status monitoring, enabling the model to autonomously identify traits signaling bearing defects, such as issues with the external raceway or lubricant degradation.

Moreover, a multi-tier diagnostic framework was proposed, incorporating deep belief networks for hierarchical classification within mechanical systems. Another notable progression was the deployment of deep neural networks for fault detection in induction motors, where sparse autoencoders were employed to independently extract features from untagged measurement data. Additional innovative approaches included the integration of Teager-Kaiser energy operators with deep belief networks for fault detection in reciprocating compressor valves, utilizing a fusion of vibrational, pressure, and current signals. Furthermore, a comprehensive review of existing literature explored the utilization of machine learning algorithms in wireless sensor networks, while other studies concentrated on multi-sensor diagnostic systems and condition-monitoring techniques based on statistical temporal features and neural networks for identifying bearing anomalies.

Another review focused on fault detection methods for electrical machines, particularly in online stator inter-turn fault diagnosis in motors. This evolution from earlier methods to more advanced, data-driven approaches marks a significant progression in the field of fault prediction and analysis for industrial systems.

In this section, we examine the key industries and equipment where the research models have been applied. Initial observations suggest that a significant portion of these applications, nearly three-quarters, are related to industrial and power systems. Other notable areas include process systems, information technology, and construction. Detailed analysis of the specific systems and equipment used to implement these models is provided in subsequent tables for each period.

In the initial phase (1989-1998), research on fault prediction was mainly concentrated on power systems, with notable applications in induction motors and transformers. Other significant areas of focus included process industries, particularly in catalytic cracking, polymerization, and cooling processes. In the subsequent phase (1999-2008), research continued to emphasize power systems, particularly components like motors, transformers, and transmission lines. However, there was a noticeable increase in the exploration of industrial systems such as turbines, compressors, and pumps. This shift reflected a growing trust in the reliability of power systems, which facilitated the emergence of new applications and an expansion of the research scope.

During the third period (2009-2018), there was a notable shift in focus towards industrial systems and their components, which made up over half of the research. While power systems still remained an area of interest, they accounted for only about a quarter of the studies, primarily centered on motors, transformers, and mechanical parts of electrical systems like bearings. Additionally, the research scope expanded to include fields such as information technology, reaching into various engineering domains beyond the traditional systems. This growing research trend underscores the increasing demand and potential benefits of these predictive models in diverse industrial applications.

Over the three periods, the research gradually evolved from addressing basic applications such as motors and transformers to tackling more complex and critical systems in industries like petrochemicals, oil and gas, and various manufacturing sectors. Additionally, critical infrastructure became the focus of attention. Research also expanded to include software applications and civil engineering structures, marking a clear shift from simpler systems to more critical and complex ones, reflecting the growing confidence in the efficacy and impact of these predictive models.

6. Research Analysis and Future Research Directions

The use of Fault prediction to improve asset reliability and optimize performance has garnered significant attention from both the academic and industrial sectors. This area has gained traction in recent years due to the advancements in artificial intelligence (AI)-driven algorithms, which address the complexities of modern processes, systems, and costly assets such as turbines, compressors, and more. Industries that rely heavily on assets, like the oil and gas sector, are increasingly seeking AI-enabled, data-driven solutions to enhance their operational efficiency and reduce losses through effective early detection and comprehensive preventive maintenance strategies. In the past, early-stage fault prediction posed significant challenges. However, with the growing availability of data and advancements in computational intelligence and machine learning (ML) technologies, the development of more sophisticated solutions has become feasible.

The interest in AI-powered solutions for fault prediction has grown considerably. Looking back at the first decade of fault prediction research, it was evident that progress was both limited in scale and scope. The research primarily relied on statistical methods, condition monitoring, and other basic tools to detect faults. These early methods often did not support real-time fault detection, leaving insufficient time for engineers to intervene and prevent unplanned downtimes. During this period, industrial applications were relatively narrow, with a focus on systems like transformers and motors. However, in the second period, fault prediction research experienced significant expansion. This era saw notable advancements in the development of more sophisticated models for fault prediction and diagnosis. As a result, the research scope widened to include new systems and processes, such as those in the oil and gas sector and other industrial applications.

The third period (the past decade) marked a period of remarkable progress. Advances in computational intelligence, data analytics, and machine learning have made it possible to predict machinery faults with much greater accuracy. During this time, the scope of fault prediction applications broadened significantly, expanding from power systems to encompass fields like information technology and a variety of industrial systems. In earlier sections of this review, we examined the development of fault prediction research over the last three decades, exploring its scope, methodologies, algorithms, and the diversity of applications that have emerged.

The connection between big data and system reliability is well-established. Big data's impact on improving reliability engineering has garnered significant attention, with recent research highlighting its value as a crucial asset, especially in industries like oil and gas. Alongside the integration of big data, several emerging research trends in machinery fault diagnosis have been proposed. These trends include shifting from behavioral studies to exploring fundamental mechanisms, transitioning from qualitative to quantitative techniques, broadening the scope from individual fault analysis to examining groups of faults, considering both major and minor faults, and progressing from component-specific fault detection to analyzing faults at the system level. In the realm of AI for diagnosing faults in rotating machinery, researchers have explored a variety of intelligent algorithms, such as k-nearest neighbors, naive Bayes, support vector machines (SVM), artificial neural networks (ANN), and deep learning methods. These techniques have shown great potential in improving the precision and effectiveness of fault diagnosis across a wide range of industrial applications.

Unplanned equipment downtime and performance issues in industries that rely heavily on assets are critical concerns. Timely fault detection is essential to minimize the impact of system failures and performance declines. A variety of techniques and models, often requiring technical expertise and relatively small amounts of data, are commonly used for fault prediction , such as knowledge-based and data-driven approaches. The increasing complexity of modern systems, combined with the massive amounts of data available and the unpredictable nature of these systems, presents significant challenges when trying to apply traditional analytical and knowledge-based models with precision. As noted in the third phase of research discussed in this review, artificial intelligence (AI), and particularly machine learning (ML), has proven to be a highly effective approach for addressing these challenges. These data-driven models are capable of managing large datasets from intricate systems, extracting meaningful insights to forecast anomalies, which is vital for industrial operations. Additionally, deep learning, a branch of AI, holds great promise in the context of big data, with the potential to create highly accurate and precise predictive models.

In machine learning applications, data plays a crucial role in empowering intelligent algorithms to identify faults and enhance the reliability of assets across industries. The Internet of Things (IoT) serves as an essential platform for gathering real-time data, which is a core element in fault prediction models used across various sectors. Together with technologies like cloud computing, IoT enables the generation, storage, and cleaning of data, making it ready for fault prediction applications. The rise of IoT has been instrumental in driving the fourth industrial revolution, or Industry 4.0, by enabling the real-time collection and analysis of equipment and manufacturing data from diverse systems. In this context, big data analytics and deep learning methods, which process both offline and live data streams,

DOI: 10.55214/25768484.v9i5.7072

are being effectively applied to fault diagnosis. This has greatly improved the performance of complex systems, leading to enhanced operational efficiency and higher profitability in a range of industries.

While there is broad agreement that the oil and gas industry stands to gain the most from fault prediction advancements through AI-driven big data models, the impact extends far beyond this sector. The use of cloud computing has the potential to greatly benefit fault prediction and systems reliability. Cloud-based solutions offer the computational power needed for real-time decision-making regarding system reliability and maintenance. Additionally, new simulation-based optimization frameworks, inspired by nature and designed for multi-criteria and multi-objective analysis, can be leveraged for a range of tasks, including predictions, diagnostics, prognostics, and more. These advancements have the potential to serve a wide array of industrial applications, enhancing their efficiency and effectiveness.

7. Conclusions

This paper conducts a systematic literature review (SLR) using bibliometric analysis (BA) tools to examine fault prediction with a focus on equipment reliability. The review is based on a decade of research papers, specifically those published between 2010 and 2020, gathered from various bibliographic databases. To assess the development of the field, we performed a co-citation network analysis, followed by a content examination of the most frequently cited papers in each period. Additionally, the top 25% of these papers were examined and classified into three main themes: the research scope, the algorithms and models employed, and the industrial applications. We enhanced the co-citation analysis by using modularity-based methods to track the evolution of research clusters over time. This paper serves as a comprehensive source for understanding the progression of fault prediction research, including the tools and models that have been applied. Despite our best efforts to encompass all relevant aspects, we acknowledge that certain references may have been missed, primarily due to the narrower time frame considered, which focused solely on the past decade. This study provides valuable insights into future directions for developing innovative methods and tools to advance fault prediction in complex systems. Our findings suggest that fault prediction is still an emerging research area. Expanding fault prediction to include related fields such as prognostics, classification, prediction, and performance analysis could enhance system reliability, optimize operational efficiency, and reduce maintenance costs by preventing breakdowns and bottlenecks. Ultimately, this paper highlights the existing progress in the field, which will drive further innovation in predictive analytics, machine learning, and smart algorithms for fault prediction applications in critical and complex systems.

Transparency:

The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

References

[1] M. Alauddin, F. Khan, S. Imtiaz, and S. Ahmed, "A bibliometric review and analysis of data-driven fault detection and diagnosis methods for process systems," *Industrial & Engineering Chemistry Research*, vol. 57, no. 32, pp. 10719–10735, 2018. https://doi.org/10.1021/acs.iecr.8b00936

[2] P. Tamilselvan and P. Wang, "Failure diagnosis using deep belief learning based health state classification," *Reliability Engineering & System Safety*, vol. 115, pp. 124–135, 2013. https://doi.org/10.1016/j.ress.2013.02.022

[3] C. Li, R. V. Sánchez, G. Zurita, M. Cerrada, D. Cabrera, and R. E. Vásquez, "Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis," *Neurocomputing*, vol. 168, pp. 119–127, 2015. https://doi.org/10.1016/j.neucom.2015.06.008

[4] K. Bacha, S. Souahlia, and M. Gossa, "Power transformer fault diagnosis based on dissolved gas analysis by support vector machine," *Electric Power Systems Research*, vol. 83, no. 1, pp. 73–79, 2012. https://doi.org/10.1016/j.epsr.2011.09.012

[5] L. Li, Y. Wang, and K.-Y. Lin, "Preventive maintenance scheduling optimization based on opportunistic production-maintenance synchronization," *Journal of Intelligent Manufacturing*, vol. 32, no. 2, pp. 545-558, 2021. https://doi.org/10.1007/s10845-020-01545-9

[6] D. Nimmo, *The political pundits* Boston: Allyn and Bacon, 1995.

[7] J. Gertler, *Fault detection and diagnosis in engineering systems*. New York: Marcel Dekker, 1998.

[8] M. A. Massoumnia, "A geometric approach to the synthesis of failure detection filters," Doctoral Dissertation, Massachusetts Institute of Technology (MIT), Cambridge, MA, 1986.

[9] M. Basseville and I. V. Nikiforov, *Detection of abrupt changes: Theory and application*. Englewood Cliffs, NJ: Prentice Hall, 1993.

[10] D. Silver *et al.*, "Mastering the game of Go without human knowledge," *Nature*, vol. 550, no. 7676, pp. 354–359, 2017. https://doi.org/10.1038/nature24270

[11] D. Silver *et al.*, "Mastering the game of Go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, 2016. https://doi.org/10.1038/nature16961

[12] K. T. Wang and W. W. Wang, "Competition in the stock market with asymmetric information," *Economic Modelling*, vol. 61, pp. 40–49, 2017. https://doi.org/10.1016/j.econmod.2016.11.024

[13] J. Zhu, L. Zheng, and K. Suyama, "Misexpression of an activated form of the FGF receptor Breathless in conjunction with downstream-of-FGF-receptor (Dof) impairs eye development in Drosophila," *Genetics*, vol. 170, no. 2, pp. 767–777, 2005.

[14] R. Isermann, "Process fault detection based on modeling and estimation methods," *Automatica*, vol. 20, no. 4, pp. 387–404, 1984. https://doi.org/10.1016/0005-1098(84)90082-0

[15] J. Chen and R. J. Patton, "Active fault-tolerant flight control systems design using the linear matrix inequality method," *Journal of Guidance, Control, and Dynamics*, vol. 22, no. 6, pp. 892–900, 1999. https://doi.org/10.2514/2.4236

[16] C. Persis and A. Isidori, "A geometric approach to nonlinear fault detection and isolation," *IEEE Transactions on Automatic Control*, vol. 46, no. 6, pp. 853–865, 2001. https://doi.org/10.1109/9.928285

[17] J. Wang and H. Hu, "Vibration-based fault diagnosis of pump using fuzzy technique," *Measurement*, vol. 39, no. 2, pp. 176–185, 2006. https://doi.org/10.1016/j.measurement.2005.07.015

[18] H. Wu, H. Jiang, and D. Shan, *Transformers fault diagnosis based on support vector machines and dissolved gas analysis. In D. S. Huang, D. C. Wunsch, D. S. Levine, & K. H. Jo (Eds.), Advanced Intelligent Computing Theories and Applications. With Aspects of Theoretical and Methodological Issues*. Springer. https://doi.org/10.1007/978-3-540-87442-3_89, 2008.

[19] J. L. Guardado, J. L. Naredo, P. Moreno, and C. R. Fuerte, "A comparative study of neural network efficiency in power transformer diagnosis using dissolved gas analysis," *IEEE Transactions on Power Delivery*, vol. 16, no. 4, pp. 643–647, 2001. https://doi.org/10.1109/61.956751

[20] A. R. G. Castro and V. Miranda, "Knowledge discovery in neural networks with application to transformer failure diagnosis," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 717–724, 2005. https://doi.org/10.1109/TPWRS.2005.846183

[21] G. Naresh, "Optimized deep learning algorithm for gearbox fault detection using vibration, sound, and acoustic emission signals," *Journal of Intelligent & Fuzzy Systems*, vol. 38, no. 3, pp. 2927–2935, 2020. https://doi.org/10.3233/JIFS-179295

[22] S. Fei and X. Zhang, "Fault diagnosis of power transformer based on support vector machine with genetic algorithm," *Expert Systems with Applications*, vol. 36, no. 8, pp. 11352–11357, 2009. https://doi.org/10.1016/j.eswa.2009.03.022

[23] A. Shintemirov, W. Tang, and Q. H. Wu, "Power transformer fault classification based on dissolved gas analysis by implementing bootstrap and genetic programming," *IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and Reviews*, vol. 39, no. 1, pp. 69–79, 2009. https://doi.org/10.1109/TSMCC.2008.2007253