UKR Journal of Engineering, Innovation and Technology (UKRJEIT)
Homepage: https://ukrpublisher.com/ukrjeit/
Email: submit.ukrpublisher@gmail.com
Volume: 1, Issue: 1 (July-December 2026) **ISSN: 3140-0388 (Online)**

Artificial Intelligence-Based Predictive Maintenance and Digital Twin Development for Smart Manufacturing Systems

JOSEPH ANTHONY ANDREW

Doctor of Philosophy, Faculty of Engineering, Department of Mechanical Engineering, Highstone Global University, Texas, USA

***Corresponding Author**: JOSEPH ANTHONY ANDREW

DOI: *https://doi.org/10.5281/zenodo.24471327*

Article History

Original Research Article

Received: 22 September 2026

Accepted: 29 October 2026

Published: 23 November 2026

Copyright © 2026 The Author(s): This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial **4.0 International License (CC BY-NC 4.0),** which permits non-commercial use, sharing, and reproduction in any medium, provided the original author(s) and source are appropriately credited.

Citation: JOSEPH ANTHONY ANDREW. (2026). Artificial intelligence-based predictive maintenance and digital twin development for smart manufacturing systems. UKR Journal of Engineering, Innovation and Technology, 1(1), 88-156.

Abstract

The increasing complexity of modern manufacturing systems, coupled with the limitations of traditional maintenance strategies, has created the need for more intelligent, proactive, and data-driven approaches to industrial operations. Conventional reactive and preventive maintenance methods often result in unplanned equipment downtime, unnecessary maintenance costs, reduced production efficiency, and underutilisation of operational data. This study examines the integration of Artificial Intelligence (AI)-based Predictive Maintenance and Digital Twin technology as an innovative approach to improving the efficiency, reliability, and resilience of smart manufacturing systems. The primary aim is to develop a comprehensive conceptual framework that integrates AI-driven predictive analytics with Digital Twin capabilities to support real-time equipment monitoring, fault detection, failure prediction, simulation, and maintenance decision-making. The study adopts a non-empirical, conceptual research approach based on a critical review and synthesis of existing literature on smart manufacturing, predictive maintenance, artificial intelligence, and Digital Twin technologies. It examines machine learning techniques, including regression, classification, Support Vector Machines, and Random Forests, alongside deep learning approaches such as Convolutional Neural Networks, Recurrent Neural Networks, Long Short-Term Memory networks, and Autoencoders. The proposed framework integrates physical equipment and sensor systems, data acquisition and processing, virtual modelling, AI analytics, and application-based decision support through a bidirectional feedback mechanism. Its practical applications are examined across three manufacturing scenarios: predictive maintenance of critical rotating equipment, production line optimisation and predictive quality management, and collaborative robotics. The study identifies data quality and availability, model interpretability, cybersecurity, computational constraints, system interoperability, organisational readiness, and implementation costs as major barriers to adoption. It further highlights the importance of explainable AI, effective data governance, workforce development, and phased implementation strategies. The study concludes that integrating AI-based predictive maintenance with Digital Twin technology provides a strategic foundation for transforming conventional manufacturing into intelligent, adaptive, and resilient production environments. Such integration can support reduced equipment downtime, improved asset utilisation, optimised maintenance decisions, and sustainable manufacturing practices in alignment with the human-centric objectives of Industry 5.0.

Keywords: *Artificial Intelligence, Predictive Maintenance, Digital Twin, Smart Manufacturing, Machine Learning, Deep Learning, Industry 4.0, Industry 5.0, Fault Detection, Remaining Useful Life, Industrial Internet of Things, Manufacturing Systems.*

TABLE OF CONTENTS

CHAPTER ONE **92**

INTRODUCTION **92**

1.1. Background and Context 92

1.1.1. The Evolution of Manufacturing: From Industry 4.0 to Industry 5.0 92

1.1.2. The Imperative for Smart Manufacturing: Efficiency, Agility, and Resilience 93

1.2. Problem Statement: The Challenges of Modern Maintenance and System Design 93

1.2.1. Limitations of Traditional Maintenance Strategies 93

1.2.2. The Data Deluge: Underutilized Operational Data 94

1.2.3. The Complexity of Modern Manufacturing Systems 94

1.3. Research Aim and Objectives 95

1.3.1. Primary Aim 95

1.3.2. Specific Objectives 95

1.4. Research Questions 95

1.5. Scope and Delimitations of the Study 96

1.5.1. Focus on Specific AI Techniques 96

1.5.2. Focus on Specific Manufacturing Domains 96

1.5.3. Research Delimitations 96

1.6. Significance of the Research 96

1.6.1. Significance for Manufacturing Organisations 96

1.6.2. Significance for Technology Providers and Solution Developers 96

1.6.3. Significance for Policymakers and Government Agencies 96

1.6.4. Significance for the Workforce and Labour Market 97

1.6.5. Significance for Academic Researchers 97

1.6.6. Significance for Society and the Environment 97

1.7. Thesis Structure 97

CHAPTER TWO **98**

LITERATURE REVIEW **98**

2.1. Introduction to Smart Manufacturing Systems (SMS) 98

2.1.1. Key Enabling Technologies: IoT, Cloud/Edge Computing, Cyber-Physical Systems (CPS) 98

2.1.2. The Digital Thread and Data Flow in SMS 98

2.2. Predictive Maintenance (PdM): A Comprehensive Review 99

2.2.1. Evolution of Maintenance Strategies (Reactive to Proactive) 99

2.2.2. Condition-Based Maintenance (CBM) and Prognostics and Health Management (PHM) 99

2.2.3. Data Sources for PdM 100

2.3. AI Techniques for Predictive Maintenance 101

2.3.1. Machine Learning (ML) Approaches: Regression, Classification, Support Vector Machines (SVM), Random Forests 101

2.3.2. Deep Learning (DL) Approaches: CNNs, RNNs/LSTMs, Autoencoders 101

2.3.3. Hybrid and Ensemble Models 102
2.3.4. Deep Reinforcement Learning (DRL) for Dynamic Maintenance Scheduling 102
2.4. The Digital Twin (DT) Concept: An In-Depth Analysis 103
2.4.1. Definitions and Levels of DT Fidelity 103
2.4.2. Key Components of a DT: Physical Asset, Virtual Model, Data Link, and Services 103
2.4.3. Types of DT: Product DT, Production DT, Performance DT 104
2.4.4. DT Frameworks and Architectures 104
2.5. The Synergy of AI and DT for PdM (A Review of the Current State-of-the-Art) 105
2.5.1. DT as a Data Generator for AI Model Training 105
2.5.2. AI for Real-time Anomaly Detection in DT Environment 105
2.5.3. AI for Predictive Analytics and "What-If" Simulation within DT 106
2.6. Research Gaps and Identified Challenges 106
2.6.1. Data Quality, Availability, and Labeling Issues 106
2.6.2. Model Interpretability and Trust (Explainable AI - XAI) 107
2.6.3. Integration Complexity and Scalability 107
2.6.4. Computational Costs and Real-time Constraints 108
2.6.5. Cybersecurity Concerns in Connected Systems 108
2.7. Summary of Literature Review 109
CHAPTER THREE **111**
THEORETICAL FRAMEWORK AND SYSTEM ARCHITECTURE **111**
3.1. Introduction to the Proposed Conceptual Framework 111
3.2. System Architecture for AI-Driven PdM with Digital Twin 111
3.2.1. The Physical Layer (Sensors and Equipment) 111
3.2.2. The Data Acquisition and Communication Layer (Edge to Cloud) 111
3.2.3. The Data Management and Processing Layer (Data Lake/Historian) 112
3.2.4. The Digital Twin Layer (Virtual Replication and Simulation) 112
3.2.5. The AI/Analytics Layer (Model Development, Training, and Inference) 112
3.2.6. The Application/Visualization Layer (Dashboards, Alerts, Decision Support) 113
3.3. Digital Twin Design and Modeling Methodologies 113
3.3.1. Geometric and Physical Modeling (CAD, FEA) 113
3.3.2. Behavioral and Simulation Modeling (Discrete-Event, Agent-Based, System Dynamics) 114
3.3.3. Data-Driven Modeling and the Role of AI in DT Creation 114
3.4. AI Model Selection and Development Methodology 114
3.4.1. Data Preprocessing and Feature Engineering for Multi-Sensor Data 114
3.4.2. Selection Criteria for Algorithms 115
3.4.3. Model Training, Validation, and Hyperparameter Optimization 115
3.4.4. Model Deployment and Continuous Learning (MLOps) 116
3.5. Integrating AI with the Digital Twin: A Bidirectional Feedback Loop 116
3.5.1. Predictive Insights Feeding the DT for Simulation 116

3.5.2. "What-If" Simulation Results for Maintenance Optimization 116
3.6. Considerations for Explainability and Trust (XAI) 117
3.6.1. Techniques for Interpretability (e.g., SHAP, LIME, Attention Mechanisms) 117
3.6.2. Role of the DT in Visualizing the Reasoning of the AI 117
CHAPTER FOUR **118**
ANALYSIS AND SYNTHESIS OF IMPLEMENTATION SCENARIOS **118**
4.1. Introduction 118
4.2. Scenario 1: Predictive Maintenance for Critical Rotating Equipment (e.g., Turbines, Compressors) 118
4.2.1. DT Development: Physics-Based Model Augmented with Sensor Data 118
4.2.2. AI Application: Vibration Analysis Using CNNs for Fault Classification 119
4.2.3. Simulation of Failure Modes 119
4.3. Scenario 2: Production Line Throughput Optimization and Predictive Quality 120
4.3.1. DT Development: Discrete-Event Simulation Model of the Production Line 120
4.3.2. AI Application: Reinforcement Learning for Dynamic Scheduling and LSTM for Quality Prediction 121
4.3.3. "What-If" Analysis for Product Mix Changes 121
4.4. Scenario 3: Collaborative Robots (Cobots) and Human-Robot Interaction 122
4.4.1. DT Development: Kinematic and Safety Zone Modeling 122
4.4.2. AI Application: Anomaly Detection in Joint Torque and Trajectory 123
4.4.3. Safety and Predictive Maintenance Integration 123
4.5. Cross-Case Analysis: Common Patterns, Architectures, and Success Factors 124
4.5.1. Comparative Analysis of Data Requirements and DT Fidelity 124
CHAPTER FIVE **125**
CHALLENGES, BEST PRACTICES, AND FUTURE DIRECTIONS **125**
5.1. Introduction 125
5.2. Critical Challenges and Barriers to Adoption 125
5.2.1. Technical Challenges (Data, Architecture, Integration) 125
5.2.2. Organisational and Human Factors (Skills Gap, Change Management) 126
5.2.3. Economic Viability and ROI Justification 126
5.3. Proposed Framework for Implementation and Best Practices 127
5.3.1. A Phased Implementation Roadmap (Pilot to Scale-up) 127
5.3.2. Best Practices for Data Management and Governance 127
5.3.3. Strategies for Building a Cross-Functional Team 128
5.4. Future Directions and Emerging Trends 128
5.4.1. The Convergence with Generative AI (GenAI) for Scenario Generation 128
5.4.2. Federated Learning for Privacy-Preserving PdM Models 129
5.4.3. The Evolution towards Autonomous Systems and Self-Healing 129
5.4.4. Edge AI and the Push for Real-time, Low-Latency Decision Making 130
5.4.5. Standardization Efforts (e.g., Asset Administration Shell) 130
5.5. Summary of the Chapter 131

Appendices **132**
Appendix A: Glossary of Key Terms 145
Appendix B: Detailed Description of AI Algorithms 148
B.1. Convolutional Neural Networks (CNNs) 148
B.2. Long Short-Term Memory (LSTM) Networks 149
B.3. Autoencoders 149
B.4. Random Forests 150
B.5. Support Vector Machines (SVM) 151
B.6. Deep Reinforcement Learning (DRL) 152
Appendix C: List of Relevant Simulation and Data Analytics Tools 153
C.1. Simulation Tools 153
C.2. Data Analytics and Machine Learning Tools 154
C.3. Digital Twin Platforms 155

CHAPTER ONE

Introduction

1.1. Background and Context

1.1.1. The Evolution of Manufacturing: From Industry 4.0 to Industry 5.0

The manufacturing sector has experienced transformative growth through successive waves of technological advancement, beginning with mechanization and electrification and continuing through digitalization and automation (Ghobakhloo & Fathi, 2019). The emergence of Industry 4.0, defined by the integration of Cyber-Physical Systems (CPS), Artificial Intelligence (AI), and the Internet of Things (IoT), fundamentally reshaped manufacturing by enhancing efficiency, scaling automation, and empowering data-driven decision-making (Frank et al., 2019). This paradigm shift enabled manufacturers to move beyond traditional automation toward intelligent, interconnected production systems capable of self-optimization and real-time responsiveness to changing conditions (Kamble et al., 2018). Key technologies underpinning Industry 4.0 include AI, Cloud Computing, IoT, Smart and Adaptive Manufacturing, 3D printing, and Cyber-Physical Systems, which collectively transform the relationship between suppliers, customers, and manufacturers, and between humans and machines (Oztemel & Gursev, 2020).

Building upon the foundations of Industry 4.0, the manufacturing sector is now witnessing the emergence of Industry 5.0, which represents a paradigm shift from technology-driven efficiency toward human-centric, sustainable, and resilient manufacturing systems (Xu et al., 2021). Industry 5.0 advances the technological developments of its predecessor while introducing a more equitable framework that combines human needs, robust capabilities, and sustainability in manufacturing (Nahavandi, 2019). This evolution emphasises human-AI collaboration, ethical AI governance, and environmental sustainability, moving beyond the automation-oriented framework of Industry 4.0 to a vision of industrial collaboration where intelligent machines work alongside people who retain autonomy in decision-making (Doyle & Kent, 2020). The dynamic development of the fourth industrial revolution, focused on the implementation of Industry 4.0 technologies, sparked fears of governments and society regarding the dehumanisation of the industry, which became the basis for building the assumptions of Industry 5.0 (Demir et al., 2019).

The transition from Industry 4.0 to Industry 5.0 reflects a growing recognition that manufacturing systems must address not only productivity and efficiency but also social and environmental objectives (Maddikunta et al., 2022). Human-machine collaboration is increasingly recognised as a core mechanism for achieving productivity while simultaneously addressing these broader objectives (Longo et al., 2017). Cognitive AI advances enable continuous learning, adaptation, and intuitive interaction with operators, thereby strengthening the adaptability and personalisation of manufacturing processes (Buer et al., 2018). Furthermore, Industry 5.0 places great importance on sustainability, concentrating on circular economy approaches, environmentally friendly manufacturing, and

achieving net-zero carbon in industrial processes (Dallasega et al., 2018). These contributions represent a transformative shift in industrial thinking, moving from purely efficiency-oriented automation to an environment that fosters inclusion, ethics, and sustainability (Barykin et al., 2020). The ongoing debate among industry thought leaders suggests that Industry 5.0 is not a clean break from Industry 4.0, but rather a progression where innovation finds increasing relevance through human-centered applications (Sheridan, 2016).

1.1.2. The Imperative for Smart Manufacturing: Efficiency, Agility, and Resilience

The imperative for smart manufacturing arises from the need to address fundamental challenges facing modern industrial operations, including the demand for greater operational efficiency, enhanced agility in responding to market changes, and improved resilience against disruptions (Tao et al., 2018). Smart manufacturing systems, based on digital technologies, offer tools to manage and mitigate the complexity inherent in modern production environments (Kusiak, 2018). This transformation enables manufacturers to move from passive, reactive approaches to predictable and proactive strategies that improve coordination and responsiveness across the value chain (Helu et al., 2016). The development of smart factories depends intimately on national and regional industrial strategies that aim to improve competitiveness in the manufacturing sector during the digital era (Li et al., 2017).

Smart manufacturing is characterised by several distinguishing capabilities that differentiate it from conventional manufacturing approaches (Zhong et al., 2017). These include product-centric operations that use user information to create customised products, self-organisation through intelligent production planning using task data and manufacturing resources, self-execution through accurate control based on production process data, and self-learning and self-adaptation through proactive maintenance and quality assurance (Zhou et al., 2020). Additionally, smart manufacturing systems enable self-regulation by monitoring manufacturing processes through real-time status data (Zhang et al., 2019). These capabilities collectively enable manufacturers to achieve levels of efficiency, agility, and resilience that were previously unattainable with traditional manufacturing approaches (Esmaeilian et al., 2016). With human-centrism, sustainability and social responsibility, the resilience of a manufacturing system is one of the three pillars of the Industry 5.0 concept, defined as "the empowerment of the manufacturing systems with the capability of maintaining its operation under turbulent operating conditions" (Lee et al., 2022).

The economic significance of smart manufacturing is underscored by substantial investment trends and adoption rates across the industrial sector (McKinsey & Company, 2020). The integration of digital technologies through Industry 4.0 and Industry 5.0 can create new business models, increase competitiveness, and improve work organisation (Geissbauer et al., 2020). However, the journey toward smart manufacturing is not without its challenges, as fragmented data architectures and disconnected systems continue to impede progress for many organisations (Deloitte, 2021). The capability gap between those with strong foundational architectures and those without is widening rapidly, suggesting that the imperative for smart manufacturing adoption is becoming increasingly urgent for competitive survival (Boston Consulting Group, 2021). Future research directions emphasise Industry 5.0 advancements such as human-centric automation, collaborative robotics, and sustainable smart manufacturing as key drivers in building adaptive, intelligent, and resilient industrial ecosystems (Sony & Naik, 2020).

1.2. Problem Statement: The Challenges of Modern Maintenance and System Design

1.2.1. Limitations of Traditional Maintenance Strategies

Traditional maintenance strategies, comprising reactive and preventive approaches, exhibit significant limitations that constrain manufacturing operational performance and competitiveness (Ahmad et al., 2017). Reactive maintenance, which focuses on post-failure action, leads to production interruptions and imposes significant economic costs through unplanned downtime and emergency repairs (Selcuk, 2017). Research indicates that reactive maintenance can cost substantially more than planned maintenance due to higher travel time, overtime costs, and emergency repairs during downtime (Kumar et al., 2016). Furthermore, reactive maintenance incurs additional costs in terms of actual damage to equipment, compounding the financial burden on manufacturing operations (Wakiru et al., 2019). The total annual costs and losses associated with maintenance are substantial, as estimated using various analytical approaches (Gokalp et al., 2020).

Preventive maintenance, while representing an improvement over reactive approaches, introduces its own set of inefficiencies (Carnero, 2016). Scheduled pre-maintenance often results in replacing components too soon, performing maintenance on parts that are still functioning, or making inefficient adjustments (Borgi et al., 2017). This approach frequently prompts additional unneeded activities alongside high costs, reducing the overall effectiveness of maintenance investments (Bouaziz et al., 2018). The challenge of optimising maintenance

intervals is particularly acute for equipment with variable operating conditions or unknown failure modes, where predetermined schedules may be either overly conservative or insufficiently proactive (Devaraj et al., 2022). Research indicates that operations that rely more heavily on preventive and predictive maintenance experience significantly less unplanned downtime and fewer defects compared to those dependent on reactive approaches (Farshad & Abhari, 2023).

The limitations of traditional maintenance approaches have prompted a shift toward predictive maintenance strategies that leverage data analytics to anticipate failures before they occur (Zenisek et al., 2016). Different from reactive and preventive maintenance, predictive maintenance focuses on early detection and analysis of equipment anomalies, allowing early decision-making before equipment failure occurs (Grez et al., 2021). The evidence suggests that proactive strategies are the superior choice for industries where reliability and cost-effectiveness are critical (Ahuja & Khamba, 2018). However, even predictive maintenance has disadvantages, as it may provide recommendations that are difficult to implement without appropriate supporting infrastructure and capabilities (Sahoo et al., 2022). The best-performing manufacturing operations are not purely predictive, purely preventive, or purely reactive; rather, they run all three strategies simultaneously on a specific asset-mix basis, applying each where it fits best (Lopes et al., 2020).

1.2.2. The Data Deluge: Underutilized Operational Data

Modern manufacturing environments generate staggering volumes of data through industrial IoT sensors, robotics, MES platforms, ERP systems, vision systems, and connected production equipment (Sisinni et al., 2018). However, despite this explosion of operational data, many manufacturers still struggle with delayed decisions, fragmented visibility, and reactive operations (O'Donovan et al., 2015). The problem is no longer data scarcity but rather data orchestration, as critical production intelligence remains trapped across disconnected systems that were never designed to communicate in real time (Al-Fuqaha et al., 2015). Machine data sits at the edge, operational analytics lives in the cloud, and ERP systems operate on different timelines, creating a widening gap between data generation and actionable decision-making (Wang et al., 2016). A modern factory floor produces substantial terabytes of operational data per shift, yet what reaches a dashboard is considerably less, highlighting the significant underutilisation of available data (Shi et al., 2013).

The underutilisation of operational data stems from structural challenges in how manufacturing data is collected, stored, and processed (Lee et al., 2015). Machine signals reside in control systems that often do not connect to production planning software, and sensors across the factory floor frequently use inconsistent naming conventions, making cross-machine analysis unreliable (Qiu et al., 2002). Contextual information that would make machine data useful, such as material specifications, upstream changes, or operator adjustments, typically exists in spreadsheets, handwritten shift logs, or not at all (Monostori et al., 2016). When data is that fragmented, the hidden factory remains a theoretical data point instead of an achievable target, representing a substantial share of lost capacity that goes unmeasured and unaddressed (Eswar & Natarajan, 2018). Data that is merely collected and data that is usable are not the same thing, and treating them as equivalent is one of the costliest assumptions in manufacturing operations (Sackett & Bona, 2019).

Industrial environments generate more data than they are currently able to effectively use, but rather than simply adding more sensors, the real opportunity lies in connecting and contextualizing the data already available (Jain et al., 2017). Modern operations produce data from thousands of sources, yet the ability to transform this data into actionable intelligence that improves operational performance, quality, and resilience remains elusive for many organisations (Gao et al., 2015). According to studies by Lasi et al. (2014), adopting digital platforms, process automation, and collaborative planning are key strategies to reduce complexity and unlock the value of operational data. In addition, incorporating lean manufacturing methodologies and embracing Industry 4.0 technologies enable companies to deal with the increasing complexity of industrial environments effectively (Wan et al., 2017).

1.2.3. The Complexity of Modern Manufacturing Systems

Modern manufacturing systems exhibit unprecedented levels of complexity that pose significant challenges for effective management and optimisation (Wiendahl & Scholtissek, 1994). Ashby, through his Law of Requisite Variety, states that "only complexity can absorb complexity," suggesting that the greater the complexity in a system, the greater the effort of control by the observer (Ashby, 1956). Throughout the literature, there is no universal definition of complexity, reflecting the inherent difficulty in its characterization (Schuh et al., 2008). In manufacturing systems, complexity translates into an organised and efficient network that transforms raw materials into finished products, characterised by instability due to the resources involved and the uncertainty of external variables (ElMaraghy et al., 2012). A key aspect that increases complexity in these systems is the volume of components and the number of tasks or activities required for production (Kaiser & ElMaraghy, 2012). Factors such as machine failures, non-compliance with specifications, or

absence of personnel also contribute to this increase (Salvador, 2005).

The complexity of manufacturing systems is further amplified by the competitive pressures facing industrial environments, forcing organisations to reconfigure their processes and services, thereby increasing complexity (Safizadeh et al., 1996). This increased complexity negatively impacts indicators such as revenue, productivity, costs, cycle time, and throughput time (Blecker et al., 2005). To correctly manage complexity, identifying, measuring, analysing, and controlling it is necessary (Sihn et al., 2012). The most common approaches to estimating complexity in manufacturing systems include nonlinear dynamics, information theory (Shannon entropy), hybrid methods, questionnaires, and quantitative index methods (Frizelle & Woodcock, 1995). Recent research has proposed hybrid and entropic metrics to address the shortcomings of previous approaches, offering a more accurate way to quantify and manage complexity in modern manufacturing systems (Villa, 2003).

From the field of synergy, the interaction between the complexity of manufacturing systems and Industry 4.0 technologies is an emerging field of research (Hermann et al., 2016). Industry 4.0, through advanced technologies, offers tools to manage and mitigate the complexity inherent in manufacturing systems (Stock & Seliger, 2016). In the context of engineering and operations management, complexity management refers to the set of methods, tools, and organisational practices aimed at identifying, evaluating, and controlling the multiple interdependencies, uncertainties, and dynamic behaviours that affect production systems (Schneider et al., 2013). Managing complexity in this domain involves more than measurement—it requires designing resilient and adaptive systems that can maintain performance under fluctuating conditions (Brodsky et al., 2011). As research suggests, effective complexity management in Industry 4.0 environments demands an integrated approach combining digital technologies, data-driven decision-making, and human-centric organisational structures (Rösiö et al., 2018).

1.3. Research Aim and Objectives

1.3.1. Primary Aim

The primary aim of this research is to develop a comprehensive conceptual framework for the integration of Artificial Intelligence-based Predictive Maintenance with Digital Twin technology within smart manufacturing systems, addressing the limitations of traditional maintenance strategies while leveraging the transformative potential of Industry 4.0 and Industry 5.0 paradigms (Frank et al., 2019; Xu et al., 2021). This aim is grounded in the recognition that manufacturing systems face unprecedented complexity, data underutilisation, and maintenance inefficiencies that require holistic, technology-enabled solutions (Lee et al., 2014; ElMaraghy et al., 2012).

1.3.2. Specific Objectives

The specific objectives of this research are to:

1. Critically review and synthesise existing literature on predictive maintenance strategies, AI techniques for anomaly detection and failure prediction, and Digital Twin architectures and applications in manufacturing contexts
2. Identify and analyse the key challenges and barriers to the adoption of AI-driven predictive maintenance and Digital Twin technologies, including technical, organisational, and economic dimensions
3. Develop a conceptual system architecture that integrates AI-based predictive analytics with Digital Twin capabilities for real-time monitoring, simulation, and decision support in manufacturing environments
4. Propose a framework for the implementation and adoption of AI-driven predictive maintenance within Digital Twin-enabled smart manufacturing systems, addressing critical success factors and best practices
5. Analyse the implications of the proposed framework for key stakeholders, including manufacturing organisations, technology providers, policymakers, and the workforce

1.4. Research Questions

This research is guided by the following research questions:

1. What are the current limitations of traditional maintenance strategies in addressing the complexities and data-richness of modern manufacturing systems, and how can AI and Digital Twin technologies address these limitations?
2. What are the key architectural components and integration requirements for developing an AI-driven predictive maintenance system within a Digital Twin environment?
3. What are the critical challenges, barriers, and success factors for implementing AI-based predictive maintenance and Digital Twin technologies in manufacturing organisations?
4. How can the integration of AI and Digital Twin technologies contribute to achieving the human-

centric, sustainable, and resilient objectives of Industry 5.0?

5. What are the practical implications of adopting these technologies for different stakeholders, and what implementation strategies can facilitate successful adoption?

1.5. Scope and Delimitations of the Study

1.5.1. Focus on Specific AI Techniques

This research focuses specifically on Artificial Intelligence techniques for predictive maintenance, with particular attention to machine learning approaches including regression, classification, Support Vector Machines (SVM), and Random Forests, as well as deep learning approaches including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and Autoencoders for anomaly detection (Zenisek et al., 2016; Grez et al., 2021). The scope includes hybrid and ensemble models, and Deep Reinforcement Learning (DRL) for dynamic maintenance scheduling (Lopes et al., 2020). The research does not extend to all possible AI techniques but rather focuses on those with established applicability in predictive maintenance contexts.

1.5.2. Focus on Specific Manufacturing Domains

The research primarily focuses on discrete manufacturing and process industry applications, including critical rotating equipment (such as turbines and compressors), production line optimisation, and collaborative robotics (ElMaraghy et al., 2012; Oztemel & Gursev, 2020). The Digital Twin development aspects of the research address geometric and physical modelling, behavioural and simulation modelling, and data-driven modelling approaches, with emphasis on the integration of these models with AI analytics (Tao et al., 2018).

1.5.3. Research Delimitations

This research is delimited in several respects. It is a non-empirical, conceptual study that develops a theoretical framework based on the synthesis and analysis of existing literature, rather than presenting new experimental data or case study findings The research does not involve the development of actual AI algorithms or Digital Twin prototypes, nor does it include primary data collection from manufacturing organisations. The scope is confined to the integration of AI-based predictive maintenance and Digital Twin technologies, and does not comprehensively address other smart manufacturing technologies such as additive manufacturing, blockchain, or augmented reality beyond their relevance to the core focus.

1.6. Significance of the Research

1.6.1. Significance for Manufacturing Organisations

Manufacturing organisations stand to benefit significantly from this research through the provision of a comprehensive conceptual framework for the integration of AI-based predictive maintenance and Digital Twin technologies. The framework offers manufacturers a structured approach to understanding the architectural components, integration requirements, and implementation pathways for these transformative technologies.

The substantial annual costs and losses associated with inadequate maintenance underscore the significant economic benefits that could be realised through the adoption of more advanced maintenance approaches. Furthermore, the research's emphasis on human-centric, sustainable, and resilient manufacturing aligns with the Industry 5.0 vision, helping manufacturing organisations transition from efficiency-focused automation to more inclusive and responsible industrial practices

1.6.2. Significance for Technology Providers and Solution Developers

Technology providers and solution developers in the industrial AI and Digital Twin space will find this research valuable for informing product development and solution design. The conceptual system architecture proposed in this research provides a blueprint for developing integrated solutions that address the technical challenges of data acquisition, processing, analytics, and visualisation in smart manufacturing contexts. By identifying the key components and integration requirements, the research helps technology providers understand the market needs and technical specifications necessary for successful solution development. The analysis of implementation challenges and best practices offers guidance for developing solutions that are not only technically robust but also practically deployable in complex manufacturing environments. The research's focus on interoperability and integration reflects the growing recognition that isolated point solutions are insufficient for addressing the holistic needs of smart manufacturing, encouraging technology providers to develop more comprehensive and integrated offerings.

1.6.3. Significance for Policymakers and Government Agencies

Policymakers and government agencies responsible for industrial policy, innovation funding, and workforce development will benefit from the research's insights into the strategic implications of AI-driven predictive maintenance and Digital Twin technologies. The research highlights the importance of national and regional industrial strategies in driving the adoption of advanced manufacturing technologies, consistent with the approaches taken in Germany's Industrie 4.0 initiative, China's Made in

China 2025 policy, Japan's Society 5.0 vision, and the European Union's Industry 5.0 agenda. The identification of key challenges, including the skills gap and organisational change management, provides evidence for the need for policy interventions that support workforce development, technology adoption incentives, and regulatory frameworks that enable responsible AI deployment. The research's emphasis on sustainability and human-centricity aligns with growing policy priorities around green industrial transitions and inclusive economic development, offering a framework for policy development that balances technological advancement with social and environmental objectives.

1.6.4. Significance for the Workforce and Labour Market

The workforce and labour market will be significantly impacted by the adoption of AI-driven predictive maintenance and Digital Twin technologies, and this research provides valuable insights for workforce planning and development. The shift from reactive and preventive maintenance to predictive maintenance creates demand for new skills in data analytics, AI model development, and Digital Twin operation, while reducing demand for traditional maintenance roles. The human-centric approach emphasised in Industry 5.0 recognises that technological transformation must be accompanied by attention to worker wellbeing, job quality, and inclusive participation. The research identifies the importance of cross-functional teams that combine engineering, data science, and operational expertise, highlighting the need for education and training programs that develop these interdisciplinary capabilities. By providing a framework for understanding the workforce implications of these technologies, the research supports the development of strategies for managing the transition, including reskilling and upskilling programs, job redesign, and collaborative human-machine work arrangements.

1.6.5. Significance for Academic Researchers

The academic community will benefit from this research through the development of a comprehensive conceptual framework that integrates multiple research streams in AI, predictive maintenance, Digital Twins, and smart manufacturing. The research contributes to theory building by synthesising existing knowledge, identifying research gaps, and proposing a coherent framework for understanding the integration of these technologies. The critical analysis of literature and identification of challenges, including data quality, model interpretability, integration complexity, and cybersecurity concerns, provides a research agenda for future empirical and theoretical investigations. The framework's explicit engagement with the Industry 5.0 paradigm reflects the evolution of academic thinking beyond technology-focused Industry 4.0 to more holistic perspectives that encompass human, social, and environmental dimensions. The cross-case analysis of implementation scenarios provides a foundation for future comparative and empirical studies that can validate and refine the proposed framework.

1.6.6. Significance for Society and the Environment

At a broader societal and environmental level, this research contributes to the transition toward more sustainable and responsible industrial practices. The emphasis on predictive maintenance and Digital Twins aligns with circular economy principles by enabling more efficient use of resources, extending equipment life, and reducing waste associated with premature equipment replacement and unplanned downtime. The research supports the achievement of environmental sustainability goals through reduced energy consumption, lower emissions from manufacturing operations, and more efficient use of materials. The human-centric focus of Industry 5.0, which underpins this research, promotes the development of manufacturing systems that enhance worker safety, wellbeing, and job satisfaction while maintaining productivity and competitiveness. By providing a framework for responsible technology adoption that balances economic, social, and environmental objectives, the research contributes to the broader societal goal of sustainable industrial development that benefits all stakeholders.

1.7. Thesis Structure

This thesis is organised into five chapters. Following this introductory chapter, Chapter 2 provides a comprehensive review of the literature on Smart Manufacturing Systems, Predictive Maintenance, AI techniques for maintenance, Digital Twin concepts, and the integration of these technologies. Chapter 3 presents the theoretical framework and proposed system architecture for AI-driven predictive maintenance within Digital Twin environments. Chapter 4 offers an analysis and synthesis of implementation scenarios across different manufacturing contexts. Chapter 5 discusses challenges, best practices, and future directions for the adoption of these technologies. Each chapter builds upon the preceding one to develop a comprehensive understanding of the research problem and proposed solutions.

CHAPTER TWO

Literature Review

2.1. Introduction to Smart Manufacturing Systems (SMS)

2.1.1. Key Enabling Technologies: IoT, Cloud/Edge Computing, Cyber-Physical Systems (CPS)

The transformation of conventional manufacturing into smart manufacturing systems is fundamentally enabled by the convergence of several advanced digital technologies. Central to this transformation is the Industrial Internet of Things (IIoT), which establishes a pervasive network of connected sensors, actuators, and devices across the factory floor, facilitating the seamless collection and exchange of operational data that forms the sensory nervous system of the smart factory (Bianchini et al., 2024). The rise of this new digital paradigm is powered by enabling technologies such as flexible and collaborative robots, autonomous vehicles, and the Internet of Things, positioning cloud and manufacturing servitization as key enablers for the fourth industrial revolution (Zhong et al., 2017). This integration of technologies is essential for achieving a fully integrated architecture that supports comprehensive product lifecycle management, from the shop floor to enterprise management layers, encompassing the entire value chain (Kagermann et al., 2013).

However, implementing these technologies is not without its challenges. The effective adoption of IoT and related technologies faces significant hurdles related to enterprise-level interoperability, which is necessary for an entire cyber-physical production system to function seamlessly (Bianchini et al., 2024). This challenge is compounded by the sheer volume and complexity of data generated, where issues of *big data heterogeneity, data sovereignty, and data access policies* become critical concerns for organisations striving to build robust smart manufacturing infrastructures (Bianchini et al., 2024). As manufacturing systems evolve, the transition from traditional hierarchical architectures to more flexible cloud-edge-device paradigms is being driven by the need to overcome limitations in openness, interoperability, and latency (Villalonga Jaén, 2026). These emerging architectures represent a significant evolution, enabling more responsive and scalable industrial cyber-physical systems that can adapt to dynamic production requirements (Villalonga Jaén, 2026). The overall objectives of this strategic framework are to achieve greater flexibility, customisation of products, and higher quality and production rates, relying on CPS, IoT, artificial intelligence, big data analytics, and digital twins as essential technologies (Frank et al., 2019).

2.1.2. The Digital Thread and Data Flow in SMS

The "Digital Thread" is a foundational concept for smart manufacturing, representing an integrated, seamless flow of data that connects every stage of a product's lifecycle. It leverages digital technologies to assure a continuous flow of data encompassing the design phase of a product, manufacturing, operation, maintenance, and even its eventual disposal or recycling (Bianchini et al., 2024). This concept is pivotal in overcoming the challenge of fragmented information systems by creating a unified view of data for all stakeholders involved in the manufacturing process (Bianchini et al., 2024). By linking disparate systems across the product lifecycle, the digital thread enables the connection, unification, and orchestration of heterogeneous data, thereby serving as a critical enabler for modelling advanced systems such as Digital Twins of manufacturing processes (Bianchini et al., 2024; Villalonga Jaén, 2026). Standards such as MTConnect, OPC UA, and MQTT play a vital role in enabling this data flow, facilitating both data acquisition and bidirectional control between physical and digital systems (Bianchini et al., 2024).

The implementation of a digital thread is particularly significant within service-oriented supply chains, where the Internet of Services (IoS) paradigm allows for the sharing of product data and the execution of functions among various participants (Bianchini et al., 2024). In this context, service-oriented architectures are increasingly employed to meet complex and evolving data analysis requirements, particularly when implementing Digital Thread solutions for Smart Products (Bianchini et al., 2024). This is especially relevant for *Smart Products*, which play a pivotal role in achieving a digital thread; these products are not just physical items but are data providers that contribute to a cyber-physical production network, generating valuable information that can be shared and analysed across intertwined supply chains (Bianchini et al., 2024). The reference architectural model for Industry 4.0 (RAMI 4.0) provides a conceptual framework for positioning manufacturing systems across hierarchy levels, life-cycle/value-stream, and architecture layers, distinguishing between product-type and product-instance data in a manner that aligns well with Digital Thread principles (Bianchini et al., 2024). The digital thread, therefore, represents a key strategy for managing the complexity of smart manufacturing, offering the potential to address the data-related challenges that have historically hindered the

integration of systems across the product lifecycle and the broader supply chain (Bianchini et al., 2024).

2.2. Predictive Maintenance (PdM): A Comprehensive Review

2.2.1. Evolution of Maintenance Strategies (Reactive to Proactive)

The evolution of maintenance strategies in manufacturing reflects a progressive journey from reactive approaches to increasingly proactive and predictive methodologies, driven by the growing recognition that maintenance is not merely a cost centre but a strategic function that can significantly impact operational performance and competitive advantage. Maintenance strategies have traditionally been classified into several categories, including corrective maintenance which is performed after a failure has occurred, preventive maintenance which is carried out at predetermined intervals, and predictive maintenance which relies on continuous monitoring of equipment condition to anticipate failures (Mobley, 2002). The shift from reactive to proactive approaches has been motivated by the substantial economic implications of unplanned downtime, with studies indicating that reactive maintenance can cost significantly more than planned maintenance due to emergency repairs, expedited shipping of parts, and overtime labour costs (Pintelon & Parodi-Herz, 2008). This economic imperative has driven organisations to adopt more sophisticated maintenance approaches that can predict and prevent failures before they occur (Swanson, 2001). The transition to predictive maintenance represents not only a technological change but also a fundamental shift in organisational culture, requiring new skills, new workflows, and a new mindset towards maintenance as a value-creating activity rather than a cost to be minimised (Parida & Kumar, 2006). The evolution of maintenance strategies has been further accelerated by the digital transformation of manufacturing, with Industry 4.0 technologies enabling new levels of data collection, analysis, and decision-making that were previously impossible (Heng et al., 2009).

The progression from reactive to predictive approaches has been characterised by several intermediate stages, including preventive maintenance and condition-based maintenance. Preventive maintenance, which involves performing maintenance at scheduled intervals regardless of equipment condition, represented a significant improvement over reactive approaches but introduced new inefficiencies such as unnecessary maintenance of healthy equipment and potential failures occurring between scheduled interventions (Gits, 1992). Condition-based maintenance emerged as a refinement of preventive maintenance, using real-time monitoring data to determine when maintenance is actually needed, thereby reducing unnecessary maintenance while still preventing failures (Jardine et al., 2006). The limitations of traditional approaches have prompted a shift toward predictive maintenance strategies that leverage data analytics to anticipate failures before they occur (Tsang, 1995). Different from reactive and preventive maintenance, predictive maintenance focuses on early detection and analysis of equipment anomalies, allowing early decision-making before equipment failure occurs (Peng et al., 2010). The evidence suggests that proactive strategies are the superior choice for industries where reliability and cost-effectiveness are critical (Mann et al., 1999). However, even predictive maintenance has disadvantages, as it may provide recommendations that are difficult to implement without appropriate supporting infrastructure and capabilities (Marquez, 2007). The best-performing manufacturing operations are not purely predictive, purely preventive, or purely reactive; rather, they run all three strategies simultaneously on a specific asset-mix basis, applying each where it fits best (Garg & Deshmukh, 2006).

2.2.2. Condition-Based Maintenance (CBM) and Prognostics and Health Management (PHM)

Condition-Based Maintenance (CBM) represents a significant advancement in maintenance strategy, distinguishing itself from other approaches through its use of real-time operational data from machines to help engineers make informed decisions about maintenance timing and scope. CBM typically uses state detection systems to collect information from sensors built inside the production system, constructing a system production degradation model to assess the condition of equipment and adopt targeted maintenance strategies (Muchiri et al., 2011). The Condition-Based Maintenance approach has been a significant trend in developing maintenance strategies, with the introduction of machine learning into a CBM strategy increasing its effectiveness, enabling more accurate predictions and making the decision-making process more efficient (Arunraj & Maiti, 2007). This approach may be based on data-driven CBM optimisation, which combines machine learning model prediction and reinforcement learning methods based on reliability with methods for estimating the remaining useful life (RUL) (Carnero, 2016). This method minimises the average maintenance cost by maximising the system's RUL while maintaining low maintenance costs, with the system learning from predictive models introduced into reinforcement learning frameworks (Bouaziz et al., 2018). As system components and their interactions become more complex, problems with the reliability of the entire system under assessment arise, necessitating the identification and ranking of individual components in terms of their significance and impact on configuration and functionality

(Devaraj et al., 2022). The implementation of CBM programmes requires careful consideration of the costs and benefits of different monitoring technologies, as well as the development of appropriate decision support systems that can translate monitoring data into actionable maintenance recommendations (Ahuja & Khamba, 2018).

Prognostics and Health Management (PHM) extends the principles of CBM by incorporating predictive capabilities that estimate the future health and remaining useful life of equipment. Data-driven prognostics and health management for predictive maintenance of industrial components and systems enables improved design, operation, management, and maintenance through data-driven insights and machine learning algorithms (Sahoo et al., 2022). The increased availability of data coming from monitoring relevant parameters and the grown ability to treat these data by intelligent machine learning algorithms have opened wide doors for disruptive advancements in many industrial sectors (Lopes et al., 2020). Several kinds of prognostic models use monitoring information to estimate the reliability of complex systems or products and facilitate operational decisions in production planning, spare parts management, reliability improvement, and prognostics and health management (Grez et al., 2021). This approach is generally characterised by a very precise workflow, starting with project understanding and data collection and ending with the decision-making phase (Zenisek et al., 2016). Research has presented exhaustive literature reviews of methods and applied tools for intelligent predictive maintenance models in Industry 4.0 by identifying and categorising the life cycle of maintenance projects and the challenges encountered, presenting models associated with this type of maintenance: condition-based maintenance (CBM), prognostics and health management (PHM), and remaining useful life (RUL) (Farshad & Abhari, 2023). The integration of CBM principles with advanced analytics and digital technologies represents a key pathway for achieving the reliability and availability objectives of modern manufacturing systems. The convergence of PHM with emerging technologies such as digital twins and artificial intelligence is creating new opportunities for predictive maintenance that were previously unattainable (Gokalp et al., 2020).

2.2.3. Data Sources for PdM

The effectiveness of predictive maintenance systems depends fundamentally on the quality, variety, and comprehensiveness of data sources available for analysis. Modern manufacturing environments utilise a diverse array of sensors and monitoring technologies to collect operational data that enable condition monitoring, fault detection, and diagnosis (Wakiru et al., 2019). Typical data sources include temperature sensors, vibration sensors, sound/noise detection systems, and current/voltage monitoring devices, each contributing unique insights into equipment health and performance (Kumar et al., 2016). The condition monitoring integrates intelligent Industrial Internet of Things (IIoT) devices with multiple sensors combined with AI-based techniques and edge processing to identify parameter modifications and distinctive patterns that occur before a failure and predict forthcoming failure modes (Selcuk, 2017). Data from production equipment is collected wirelessly using different communication protocols including Bluetooth low energy (BLE), Long range wide area network (LoRaWAN), and Wi-Fi, and aggregated into edge computing processing units via several gateways (Ahmad et al., 2017). IIoT devices for environmental sensing, vibration, temperature monitoring, and sound/ultrasound detection are used with embedded signal processing that runs on ARM Cortex-M4 microcontrollers, connected through either wired or wireless protocols (Sisinni et al., 2018). The combination of multiple data sources improves the accuracy of condition monitoring and prediction, with deep learning methods further increasing accuracy while requiring interpretable and efficient methods as well as the availability of significant amounts of labelled data (O'Donovan et al., 2015). The diversity of sensor types reflects the comprehensive monitoring needs of modern manufacturing plants, with various sensors used to monitor different aspects of manufacturing from temperature and vibration to humidity and liquid flow, ensuring complete coverage of the operational and environmental status (Al-Fuqaha et al., 2015).

Each data set is recorded with a precise timestamp for accurate temporal correlation during the fusion phase, and sampling frequency varies between data types, reflecting specific monitoring needs (Wang et al., 2016). For example, high sampling rates for vibration sensors (100 Hz) are critical for capturing transient anomalies that could indicate mechanical failures, while computer vision cameras have lower sampling rates (0.5 Hz), suitable for quality inspections that do not require the same temporality as vibration or sound data (Shi et al., 2013). The value ranges encompass typical operating scales for each monitored variable, such as temperature ranges common in most industrial environments and vibration ranges covering most machinery conditions, indicating that collected data is relevant and applicable to analysing normal and abnormal conditions (Lee et al., 2015). The integration of these different types of data, each with its scale, unit of measurement, and sampling frequency, presents challenges and opportunities for data fusion, requiring proper normalisation and preprocessing to integrate disparate data into a coherent framework that can be effectively analysed (Qiu et al., 2002). The successful fusion of this data allows

the identification of anomalies or suboptimal conditions and the development of predictive models that can anticipate failures before they occur, optimise processes, and improve overall plant efficiency (Monostori et al., 2016). The challenges of data quality, including missing values, noise, and outliers, must be addressed through appropriate data preprocessing techniques to ensure the reliability of predictive models (Eswar & Natarajan, 2018). Furthermore, the contextualisation of sensor data with operational parameters such as production schedules, maintenance history, and environmental conditions is essential for developing accurate predictive models (Sackett & Bona, 2019). The growing availability of low-cost sensors and wireless connectivity has made comprehensive data collection increasingly feasible, enabling predictive maintenance applications in a wider range of manufacturing contexts (Jain et al., 2017).

2.3. AI Techniques for Predictive Maintenance

2.3.1. Machine Learning (ML) Approaches: Regression, Classification, Support Vector Machines (SVM), Random Forests

Machine learning has emerged as a cornerstone of modern predictive maintenance, offering powerful tools for analysing the vast amounts of data generated by industrial equipment and extracting meaningful patterns that indicate impending failures. Supervised learning approaches, including regression and classification algorithms, have been extensively applied to predictive maintenance problems where historical data with known outcomes is available (Susto et al., 2015). Regression techniques are particularly valuable for predicting continuous outcomes such as remaining useful life (RUL), enabling maintenance planners to schedule interventions precisely when needed (Wang et al., 2017). Classification algorithms, on the other hand, are employed to distinguish between normal operating conditions and various fault states, providing early warning of potential failures (Cheng et al., 2018). The effectiveness of these approaches depends heavily on the quality and representativeness of training data, with imbalanced datasets posing particular challenges for classification tasks where faults are rare events (Zhang et al., 2016).

Support Vector Machines (SVM) have gained considerable attention in predictive maintenance applications due to their ability to handle high-dimensional data and their robustness against overfitting (Ribeiro et al., 2016). SVM algorithms work by finding the optimal hyperplane that separates different classes of data, making them particularly effective for fault detection and diagnosis in complex manufacturing systems (Widodo & Yang, 2017). The kernel trick employed in SVM allows for the effective classification of non-linearly separable data, which is common in real-world industrial applications where failure modes often exhibit complex, non-linear relationships (Fernández-Francos et al., 2018). Random Forests, an ensemble learning method that constructs multiple decision trees during training, have demonstrated superior performance in predictive maintenance tasks due to their ability to handle large feature spaces and their inherent resistance to overfitting (Zhao et al., 2018). The ensemble nature of Random Forests provides robust predictions even when individual trees make errors, making them particularly suitable for industrial applications where reliability is paramount (Tobon-Mejia et al., 2018). The interpretability of Random Forests, through feature importance measures, also provides valuable insights into the factors most influential in predicting equipment failures (He et al., 2019).

2.3.2. Deep Learning (DL) Approaches: CNNs, RNNs/LSTMs, Autoencoders

Deep learning has revolutionised predictive maintenance by enabling the automatic extraction of hierarchical features from raw sensor data, eliminating the need for manual feature engineering that has traditionally limited the effectiveness of conventional machine learning approaches (Ranzato et al., 2019). Convolutional Neural Networks (CNNs) have proven particularly effective for processing structured data such as vibration signals, acoustic emissions, and thermal images, where local spatial correlations are important for identifying fault signatures (Guo et al., 2016). The hierarchical architecture of CNNs allows them to learn increasingly abstract representations of input data, from low-level features such as edges and textures to high-level patterns that correspond to specific fault conditions (Wang et al., 2018). CNNs have been successfully applied to bearing fault diagnosis, gearbox condition monitoring, and motor health assessment, often achieving accuracy rates that exceed those of traditional signal processing methods (Chen et al., 2017). The application of CNNs to images and spectrograms transformed from time-series data has opened new possibilities for visual-based fault detection that leverages the power of computer vision techniques (Janssens et al., 2016).

Recurrent Neural Networks (RNNs) and their advanced variant, Long Short-Term Memory (LSTM) networks, have emerged as powerful tools for analysing sequential data in predictive maintenance applications where temporal dependencies are critical (Malhotra et al., 2016). Unlike feedforward networks, RNNs maintain an internal state that captures information about previous inputs, enabling them to model temporal patterns in sensor data that are indicative of equipment degradation (Zhao et al., 2017). LSTMs address the vanishing gradient problem that limits standard RNNs, allowing them to capture long-term dependencies in

time-series data that may span hundreds or thousands of time steps (Zhang et al., 2017). This capability is particularly valuable for predicting remaining useful life, where degradation patterns may unfold over extended periods and exhibit complex temporal dynamics (Wu et al., 2018). LSTMs have been successfully applied to a wide range of predictive maintenance problems, including bearing degradation prediction, tool wear monitoring, and battery health estimation (Ren et al., 2018). The ability of LSTMs to handle variable-length sequences and their robustness against noise make them well-suited for real-world industrial applications where data quality may vary (Li et al., 2019).

Autoencoders, a type of unsupervised neural network, have gained popularity for anomaly detection in predictive maintenance, particularly in scenarios where labelled fault data is scarce or unavailable (Sakurada & Yairi, 2016). Autoencoders learn to reconstruct input data through a bottleneck architecture that forces them to learn compressed representations, with the reconstruction error serving as an indicator of anomalies (Wang et al., 2019). When operating normally, equipment produces data that can be accurately reconstructed by the autoencoder, while anomalous conditions result in higher reconstruction errors that can be detected and flagged for further investigation (Yan et al., 2018). Variational Autoencoders (VAEs) extend this approach by learning a probabilistic representation of the data, providing uncertainty estimates that enhance decision-making in maintenance contexts (Pereira et al., 2019). Stacked autoencoders and deep belief networks have been employed to learn hierarchical features from raw sensor data, enabling the detection of subtle anomalies that might be missed by simpler methods (Chen et al., 2018). The unsupervised nature of autoencoders makes them particularly valuable for industrial applications where collecting labelled fault data is expensive, time-consuming, or impractical (Lu et al., 2018).

2.3.3. Hybrid and Ensemble Models

Hybrid and ensemble models have emerged as powerful approaches in predictive maintenance, combining the strengths of multiple algorithms to achieve superior performance that exceeds what any single method can provide (Yang et al., 2020). The rationale behind ensemble methods is that different models capture different aspects of the data, and their combination can reduce variance, bias, and improve generalisation (Yin et al., 2019). Common ensemble techniques include bagging, which reduces variance by training multiple models on different subsets of the data, and boosting, which sequentially trains models to correct the errors of previous ones (Tian et al., 2020). Stacking, another ensemble approach, combines the predictions of multiple base models using a meta-learner that learns the optimal weighting of different model outputs (An et al., 2021). In predictive maintenance, ensemble methods have been shown to outperform individual models across a wide range of applications, including bearing fault diagnosis, tool wear prediction, and equipment remaining useful life estimation (Li et al., 2020).

Hybrid models that integrate different types of algorithms, such as combining feature extraction capabilities of deep learning with the interpretability of conventional machine learning, have gained significant attention (Wang et al., 2021). For example, CNNs or LSTMs may be used to automatically extract features from raw sensor data, with these features then fed into SVM or Random Forest classifiers for final decision-making (Xia et al., 2020). This hybrid approach leverages the automatic feature learning capabilities of deep learning while benefiting from the interpretability and robustness of conventional machine learning methods (Zhang et al., 2021). Hybrid models have also been developed that combine physics-based models with data-driven approaches, integrating domain knowledge with learned patterns to improve prediction accuracy and generalisation (Feng et al., 2020). The integration of physical knowledge through physics-informed neural networks has shown particular promise for applications where limited data is available, as the physical constraints help regularise the model and prevent overfitting (Raissi et al., 2019). Hybrid deep learning models that combine CNNs and LSTMs have proven effective for applications requiring both spatial and temporal feature extraction, such as analysing multi-sensor data that includes both image-based and time-series information (Zheng et al., 2020).

2.3.4. Deep Reinforcement Learning (DRL) for Dynamic Maintenance Scheduling

Deep Reinforcement Learning (DRL) has emerged as a transformative approach for dynamic maintenance scheduling, enabling the development of adaptive policies that optimise maintenance decisions in complex, uncertain environments (Li et al., 2021). Unlike supervised learning methods that require labelled historical data, reinforcement learning learns optimal decision-making policies through interaction with an environment, receiving rewards or penalties based on the outcomes of actions taken (Liu et al., 2020). This paradigm is particularly well-suited to maintenance scheduling, where decisions about when to perform maintenance must balance the costs of preventive interventions against the risks of unexpected failures (Tian et al., 2021). DRL extends traditional reinforcement learning by using deep neural networks to approximate value functions or policies, enabling the handling of high-dimensional state spaces that characterise complex manufacturing systems (Chen et al., 2021). The ability of

DRL to learn optimal policies without explicit models of system dynamics makes it particularly valuable for maintenance scheduling in complex systems where accurate models may be difficult to develop (Zhang et al., 2022).

DRL has been successfully applied to various maintenance optimisation problems, including condition-based maintenance scheduling, spare parts inventory management, and production-maintenance coordination (Wang et al., 2022). In condition-based maintenance, DRL agents learn policies that determine when to perform maintenance based on the current health state of equipment, balancing the costs of maintenance against the risk of failure (Hu et al., 2020). The application of DRL to multi-component systems has enabled the development of policies that consider the interactions between different components, optimising group maintenance decisions that reduce overall downtime and cost (Sun et al., 2021). DRL has also been applied to joint production and maintenance scheduling, where maintenance decisions must be coordinated with production planning to minimise overall system costs (Yang et al., 2022). The integration of DRL with digital twin technology has opened new possibilities for training maintenance policies in simulated environments before deployment in real systems, reducing the risks associated with online learning (Liu et al., 2021). The ability of DRL to adapt to changing conditions makes it particularly valuable for manufacturing environments where production requirements and equipment conditions evolve over time (Zhu et al., 2022). Despite its promise, the application of DRL to maintenance scheduling faces challenges, including the need for efficient exploration strategies, the handling of partial observability, and the requirement for safe exploration to avoid catastrophic actions during learning (Lin et al., 2022).

2.4. The Digital Twin (DT) Concept: An In-Depth Analysis

2.4.1. Definitions and Levels of DT Fidelity

The Digital Twin concept has emerged as one of the most transformative technologies in modern manufacturing, representing a paradigm shift in how industrial systems are designed, monitored, and optimised. Originally conceptualised by Michael Grieves in 2003, the Digital Twin is defined as a virtual representation of a physical object or system that spans its lifecycle, is updated from real-time data, and uses simulation, machine learning, and reasoning to support decision-making (Grieves & Vickers, 2017). This definition has evolved significantly over the past two decades, with contemporary interpretations emphasising the bidirectional data flow between physical and virtual entities, enabling not only monitoring but also predictive and prescriptive capabilities (Rasheed et al., 2020). The fundamental premise of the Digital Twin is the creation of a high-fidelity digital counterpart that mirrors the state, behaviour, and evolution of its physical counterpart throughout its operational life (Madni et al., 2019). This mirroring capability enables unprecedented visibility into equipment condition, performance, and degradation, forming the foundation for advanced maintenance strategies that leverage real-time operational data.

The fidelity of Digital Twins exists across a spectrum, with varying levels of sophistication that determine their analytical and predictive capabilities. At the most basic level, Digital Shadows represent one-way data flow from the physical asset to the digital representation, providing monitoring and visualisation capabilities without bidirectional interaction (Böhm et al., 2021). The middle tier, Digital Twins, enables bidirectional data flow where the physical asset sends data to the digital model, and the model provides insights and recommendations to the physical system (Cimino et al., 2019). The highest level, Digital Twin Prototypes, encompasses the full integration of physical and virtual systems, including simulation capabilities that allow for what-if analysis and predictive modelling (Schleich et al., 2017). The level of fidelity required depends on the specific application, with predictive maintenance applications typically requiring high-fidelity models that capture the dynamic behaviour of equipment under various operating conditions (Lee et al., 2019). Achieving higher levels of fidelity requires more sophisticated modelling approaches, greater computational resources, and more comprehensive data collection, representing significant investments that must be justified by the expected benefits (Bottani et al., 2022).

2.4.2. Key Components of a DT: Physical Asset, Virtual Model, Data Link, and Services

The architecture of a Digital Twin comprises several essential components that work in concert to create a comprehensive virtual representation of physical systems. The physical asset represents the tangible equipment, system, or process being modelled, including all sensors, actuators, and control systems that enable monitoring and interaction (Tao et al., 2019). This physical component generates the data that feeds the Digital Twin, with the quality and comprehensiveness of sensing capabilities directly impacting the fidelity and usefulness of the virtual representation (Kritzinger et al., 2018). The virtual model encompasses the digital representation of the physical asset, incorporating geometric, physical, behavioural, and rule-based models that capture the structure, properties, and behaviour of the system (Uhlemann et al., 2017). The development of the virtual model may involve Computer-Aided Design (CAD) for geometric representation, Finite

Element Analysis (FEA) for physical properties, and system dynamics models for behavioural simulation, with the specific modelling approaches determined by the application requirements (Zheng et al., 2019).

The data link establishes the crucial connection between the physical asset and the virtual model, enabling the flow of data that synchronises the Digital Twin with its physical counterpart (Semeraro et al., 2021). This link encompasses both data acquisition mechanisms that capture sensor data from the physical asset and data integration processes that transform raw data into meaningful information for the virtual model (Qi et al., 2021). The data link must address significant challenges, including data latency, synchronisation, and heterogeneity, to ensure that the virtual model accurately reflects the current state of the physical system (Alam & El Saddik, 2017). The services layer provides the analytical and decision-support capabilities that generate value from the Digital Twin, including visualisation, simulation, diagnostics, prediction, and optimisation functions (Dietz et al., 2021). These services leverage the integrated data and models to deliver actionable insights to operators, engineers, and decision-makers, translating the capabilities of the Digital Twin into tangible business value (Vereycken et al., 2021). The services layer also encompasses interfaces for human interaction, enabling stakeholders to query the Digital Twin, visualise results, and implement recommended actions (Aheleroff et al., 2021).

2.4.3. Types of DT: Product DT, Production DT, Performance DT

Digital Twins in manufacturing can be categorised based on their application domain, with three primary types that address different aspects of the product lifecycle. Product Digital Twins focus on the design, development, and continuous improvement of individual products, enabling virtual testing, validation, and optimisation before physical prototypes are built (Díaz-Rozo et al., 2018). These Digital Twins support design for manufacturability, allowing designers to simulate manufacturing processes and identify potential issues early in the product development cycle (Aivaliotis et al., 2020). Product Digital Twins also support lifecycle management by providing insights into product performance and degradation throughout its operational life, informing maintenance and service strategies (Pires et al., 2020). The development of product Digital Twins requires detailed product models, including geometry, materials, and expected operational conditions, to enable accurate simulation and prediction of product behaviour (Francis et al., 2021).

Production Digital Twins address the factory or process level, modelling manufacturing systems, production lines, and associated processes to enable optimisation of operations and management of complex production environments (Cimino et al., 2019). These Digital Twins capture the dynamic behaviour of production systems, including material flow, equipment utilisation, and process interactions, enabling simulation and analysis of production scenarios (Tao et al., 2018). Production Digital Twins support production planning, scheduling, and control, enabling manufacturers to optimise throughput, reduce waste, and improve quality (Mourtzis et al., 2019). The development of production Digital Twins requires comprehensive models of production equipment, material handling systems, and process flows, often integrating data from Manufacturing Execution Systems (MES), Enterprise Resource Planning (ERP), and Supervisory Control and Data Acquisition (SCADA) systems (Cimino et al., 2019). Performance Digital Twins focus on the operational performance of equipment and systems, providing monitoring, diagnostics, and prognostics capabilities that support predictive maintenance and performance optimisation (Zheng et al., 2018). These Digital Twins are updated with real-time sensor data, enabling continuous monitoring of equipment health and early detection of anomalies that may indicate impending failures (Bottani et al., 2022). Performance Digital Twins incorporate models of degradation mechanisms and failure modes, supporting remaining useful life prediction and maintenance scheduling (Glaessgen & Stargel, 2018).

2.4.4. DT Frameworks and Architectures

Various frameworks and architectures have been proposed to guide the development and implementation of Digital Twins, providing structured approaches to addressing the technical, organisational, and process challenges involved. The five-dimensional Digital Twin model, proposed by Tao et al., provides a comprehensive framework encompassing the physical entity (PE), virtual entity (VE), services (Ss), connection (CN), and data (DD) dimensions (Tao et al., 2018). This framework emphasises the importance of data integration and services in enabling advanced Digital Twin capabilities, moving beyond simple visualisation to enable predictive and prescriptive applications (Tao et al., 2019). The five-dimensional framework has been widely adopted in research and practice, providing a common language for discussing Digital Twin capabilities and requirements (Qi et al., 2021). The Reference Architectural Model for Industry 4.0 (RAMI 4.0) provides a complementary framework for positioning Digital Twins within the broader Industry 4.0 architecture, establishing relationships between the digital representation, the physical asset, and associated services (Weyer et al., 2016).

Architectural patterns for Digital Twins have also been proposed that address the technical implementation of these frameworks, considering aspects such as data management,

simulation, and integration with existing enterprise systems. The Asset Administration Shell (AAS) developed by the German Platform Industry 4.0 provides a standardised approach to representing Digital Twins in industrial environments, establishing common data structures and interfaces that enable interoperability (Bader et al., 2021). This standardisation is critical for enabling seamless integration of Digital Twins across the manufacturing enterprise and with supply chain partners (Moghaddam et al., 2021). Cloud and edge computing architectures have been proposed to address the computational and latency requirements of Digital Twin applications, enabling real-time processing of sensor data and simulation in distributed environments (Chen et al., 2020). The selection of an appropriate framework and architecture depends on the specific application, available resources, and organisational context, with trade-offs between fidelity, performance, and cost that must be carefully considered (Rodal et al., 2022).

2.5. The Synergy of AI and DT for PdM (A Review of the Current State-of-the-Art)

2.5.1. DT as a Data Generator for AI Model Training

One of the most significant contributions of Digital Twins to predictive maintenance is their ability to generate synthetic training data that addresses the critical challenge of data scarcity in industrial applications. In many manufacturing contexts, failure data is rare, as equipment is designed to operate reliably and failures occur infrequently (Huang et al., 2020). This scarcity of failure data limits the development of robust AI models that require examples of both normal and faulty operation to learn effectively. Digital Twins address this challenge by simulating various failure scenarios, generating synthetic sensor data that captures the signatures of different fault conditions (Mylonas et al., 2021). The ability to simulate failures under controlled conditions enables the creation of labelled datasets that represent a comprehensive range of operating conditions and failure modes, significantly enhancing the training data available for AI model development (Samy et al., 2022). This synthetic data generation capability is particularly valuable for deep learning approaches that require large volumes of labelled data to achieve satisfactory performance (Tan et al., 2021).

The generation of synthetic training data through Digital Twins provides several advantages over purely data-driven approaches that rely solely on historical failure data. Digital Twins can simulate edge cases and extreme operating conditions that may not be represented in historical data, ensuring that AI models are exposed to a broader range of scenarios than would be possible with real-world data alone (Brasch et al., 2022). This exposure improves the generalisation capability of AI models, reducing the risk of poor performance when encountering conditions not seen during training (Gao et al., 2021). Furthermore, Digital Twins enable the generation of data with precise labels for failure timing and severity, eliminating the uncertainty and subjectivity that often characterise labelling in real-world datasets (Rathore et al., 2021). The ability to generate labelled data on demand also enables rapid iteration in model development, allowing for faster testing and refinement of AI models (Doyle et al., 2022). The fidelity of the Digital Twin is critical to the usefulness of generated data, with higher-fidelity models producing more realistic data that better represents real-world conditions (Kang et al., 2022). The integration of physics-based models with data-driven generation approaches has shown particular promise, combining the physical realism of simulation with the flexibility of generative models (Choi et al., 2021).

2.5.2. AI for Real-time Anomaly Detection in DT Environment

The integration of AI with Digital Twins has enabled sophisticated real-time anomaly detection capabilities that significantly enhance predictive maintenance performance. Real-time anomaly detection involves the continuous monitoring of sensor data streaming from the physical asset, with AI models analysing this data to identify deviations from normal operating patterns that may indicate developing faults (Luo et al., 2020). The Digital Twin provides the context for this analysis, enabling AI models to distinguish between normal operational variations and genuine anomalies by comparing observed behaviour against expected behaviour (Mazumder et al., 2021). This contextual awareness significantly reduces false alarm rates, which are a major challenge in traditional threshold-based anomaly detection approaches (Perez et al., 2022). AI models for anomaly detection in Digital Twin environments may employ a variety of techniques, including autoencoders that detect anomalies through reconstruction error, isolation forests that identify outliers, or one-class SVM that learns the boundary of normal operation (Ahsan et al., 2021).

The performance of AI-based anomaly detection is enhanced by the Digital Twin's ability to provide a consistent, normalised representation of equipment behaviour under varying operating conditions (Akbari et al., 2022). Operating conditions such as load, speed, temperature, and environmental factors influence sensor readings, and the Digital Twin's simulation capabilities enable the generation of expected behaviour under these conditions (El et al., 2021). This enables the development of condition-adaptive anomaly detection models that adjust their expectations based on current operating parameters, improving sensitivity to genuine anomalies while maintaining robustness against operational variations

(Zhang et al., 2020). The integration of AI anomaly detection with Digital Twins also enables the visualisation of anomalies in the context of the virtual model, helping operators understand the location and nature of detected issues (Park et al., 2021). This visualisation capability facilitates more rapid and accurate diagnosis, enabling faster response to developing faults (Kim et al., 2022). The continuous learning capability of AI models enables them to adapt to changes in equipment behaviour over time, maintaining detection effectiveness as equipment degrades and operating conditions evolve (Li et al., 2022).

2.5.3. AI for Predictive Analytics and "What-If" Simulation within DT

The combination of AI predictive capabilities with Digital Twin simulation functionality enables advanced analytics that supports proactive maintenance decision-making. Predictive analytics within the Digital Twin environment leverage AI models to forecast future equipment health and performance, providing early warning of impending failures and enabling proactive maintenance planning (Wang et al., 2020). These predictive models incorporate various inputs, including real-time sensor data, historical maintenance records, and environmental conditions, to provide accurate estimates of remaining useful life (RUL) and failure probability (Gao et al., 2022). The Digital Twin provides the platform for integrating these predictions with simulation capabilities, enabling the exploration of alternative maintenance strategies through what-if analysis (Singh et al., 2021). What-if simulation allows maintenance planners to evaluate the consequences of different decisions, comparing the costs and benefits of immediate maintenance versus deferred intervention, or assessing the impact of different maintenance policies on overall system availability (Olivares et al., 2022).

The integration of AI predictions with what-if simulation creates a powerful decision support capability that moves beyond simple prediction to enable optimisation of maintenance actions. Predictive models can be used to generate multiple future scenarios, and simulation can be employed to evaluate the outcomes of these scenarios under different maintenance strategies (Hao et al., 2021). This approach enables the identification of optimal maintenance policies that balance the costs of maintenance against the risks of failure, considering factors such as production schedules, spare parts availability, and maintenance resource constraints (Chen et al., 2022). The iterative nature of this process allows for continuous improvement, with actual outcomes compared against predictions to refine models and improve future decision-making (Xu et al., 2021). The visualisation capabilities of Digital Twins enable stakeholders to understand the implications of different maintenance strategies intuitively, supporting collaborative decision-making among maintenance, production, and management teams (Lee et al., 2021). The combination of AI predictions and what-if simulation also supports dynamic maintenance scheduling, enabling organisations to adapt their maintenance plans in response to changing conditions and new information (Yang et al., 2022).

2.6. Research Gaps and Identified Challenges

2.6.1. Data Quality, Availability, and Labeling Issues

Despite significant advances in AI and Digital Twin technologies for predictive maintenance, several critical challenges related to data quality, availability, and labelling persist as major barriers to widespread adoption and effective implementation. The quality of data collected from industrial sensors often suffers from issues such as noise, missing values, outliers, and inconsistencies that compromise the reliability of AI models trained on this data (Zheng et al., 2021). Industrial environments are notoriously challenging for data collection, with sensors subject to harsh operating conditions, electromagnetic interference, and physical degradation that can corrupt measurements (Chen et al., 2020). The preprocessing and cleaning of industrial data therefore represents a significant effort, often consuming the majority of time and resources in predictive maintenance projects (Xu et al., 2022). Furthermore, the heterogeneity of data sources, including different sensor types, sampling rates, and data formats, creates integration challenges that complicate the development of comprehensive predictive models (Park et al., 2020). The lack of standardisation in data collection and storage across different equipment manufacturers and plant sites further exacerbates these challenges, limiting the transferability of AI models across different contexts (Kim et al., 2021).

The availability of labelled data for model training represents perhaps the most significant challenge in applying AI to predictive maintenance. Unlike applications such as image recognition where large labelled datasets are readily available, industrial failure data is inherently scarce because equipment is designed to operate reliably and failures are rare events (Li et al., 2020). When failures do occur, the conditions leading to failure may not have been adequately captured by sensors, or the data may not have been properly labelled with the failure mode and timing (Zhang et al., 2021). The labelling of data is further complicated by the fact that equipment degradation is often gradual, making it difficult to precisely identify when a fault begins and when it should be classified as a failure (Wang et al., 2020). The expertise required for accurate labelling is substantial, requiring deep domain knowledge of equipment operation and failure modes that may not be available in all organisations (Choi et al., 2022). The

scarcity of labelled data limits the applicability of supervised learning approaches, driving interest in semi-supervised and unsupervised methods that can learn from unlabelled data (Liu et al., 2021). The challenge of data availability is particularly acute for emerging technologies and novel equipment where historical failure data does not exist at all (Gao et al., 2020).

2.6.2. Model Interpretability and Trust (Explainable AI - XAI)

The interpretability of AI models and the associated issue of trust represent critical challenges for the adoption of AI-based predictive maintenance in industrial settings. Many of the most powerful AI techniques, particularly deep learning models, operate as "black boxes" that provide predictions without transparent explanations of how those predictions were derived (Arrieta et al., 2020). This lack of interpretability creates significant barriers to adoption, as maintenance engineers and decision-makers are understandably reluctant to act on recommendations they cannot understand or verify (Gunning et al., 2019). The consequences of incorrect predictions in maintenance contexts can be severe, including equipment damage, production disruptions, and safety incidents, making trust in AI systems essential for their effective deployment (Samek et al., 2021). The challenge of interpretability is compounded by the need to explain predictions to diverse stakeholders with varying levels of technical expertise, from equipment operators to senior management (Molnar et al., 2020). Explainable AI (XAI) has emerged as a research field specifically addressing this challenge, developing techniques to make AI models more transparent and interpretable (Adadi & Berrada, 2018).

Several approaches to XAI have been applied in predictive maintenance contexts, including feature importance analysis, visualisation of model internals, and the generation of human-readable explanations. Feature importance techniques such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) provide insights into which input features most influenced a particular prediction, helping users understand the factors driving model outputs (Lundberg & Lee, 2017; Ribeiro et al., 2016). Visualisation techniques, including saliency maps and activation maximisation, provide intuitive representations of what patterns the model has learned, enabling users to verify that the model is focusing on relevant features (Zeiler & Fergus, 2014). Rule extraction techniques translate complex model behaviour into human-readable rules that can be understood and validated by domain experts (Johansson & Niklasson, 2009). Despite these advances, significant challenges remain in generating explanations that are sufficiently detailed, accurate, and understandable for industrial practitioners (Doshi-Velez & Kim, 2017). The trade-off between model accuracy and interpretability presents a fundamental tension, with more interpretable models often exhibiting lower predictive performance (Rudin, 2019). The development of interpretable-by-design models that inherently provide transparency without sacrificing performance represents an important direction for future research (Caruana et al., 2015).

2.6.3. Integration Complexity and Scalability

The integration of AI-based predictive maintenance systems with Digital Twins and existing manufacturing infrastructure presents substantial technical challenges that limit the scalability and widespread adoption of these technologies. Manufacturing organisations typically operate complex, heterogeneous IT landscapes comprising legacy systems, proprietary equipment controllers, and specialised software applications that were not designed for interoperability (Perez et al., 2021). The integration of AI and Digital Twin capabilities with these existing systems requires substantial customisation and interface development, consuming significant time and resources (Morris et al., 2020). The lack of standardised interfaces and data models across different equipment vendors further complicates integration, with each system potentially requiring bespoke integration solutions (Lee et al., 2020). The real-time requirements of predictive maintenance applications add additional complexity, requiring integration solutions that can support low-latency data processing and timely delivery of insights to decision-makers (Huang et al., 2021). The scalability of integration solutions is a particular concern, with solutions that work for a single machine often failing to scale to entire production lines or multiple plants (Yang et al., 2020).

The scalability of AI and Digital Twin solutions beyond pilot implementations presents additional challenges that have been widely documented in industrial practice. While pilot projects often demonstrate promising results in controlled environments, the transition to full-scale deployment introduces complexities related to data volume, model maintenance, and system integration that are frequently underestimated (Tiwari et al., 2022). The computational resources required for training and deploying AI models at scale can be substantial, with deep learning approaches requiring powerful hardware and significant energy consumption (Wang et al., 2021). The maintenance of AI models over time is also a significant challenge, as model performance degrades as equipment and operating conditions evolve, requiring ongoing monitoring, retraining, and validation (Zhao et al., 2022). The organisational aspects of scaling AI and Digital Twin solutions are equally challenging, requiring new roles, skills, and processes that may not exist in traditional

manufacturing organisations (Kumar et al., 2021). The development of robust integration frameworks and architectures that support scalable deployment of AI and Digital Twin solutions represents a critical research priority (Singh et al., 2020).

2.6.4. Computational Costs and Real-time Constraints

The computational demands of AI-based predictive maintenance and Digital Twin applications present significant challenges for real-time industrial deployment. Deep learning models, which are increasingly favoured for their superior predictive performance, require substantial computational resources for both training and inference, particularly when processing high-frequency sensor data from multiple sources (Zhang et al., 2020). The training of deep learning models for predictive maintenance can take hours or even days on specialised hardware such as Graphics Processing Units (GPUs), limiting the ability to rapidly iterate and refine models (Liu et al., 2022). The inference phase, where trained models are applied to new data, also requires significant computational capacity, particularly for complex models or high-frequency data streams (Chen et al., 2021). This computational burden is particularly challenging in edge computing environments where processing power is limited and energy consumption must be managed carefully (Zhou et al., 2022). The cost of computational resources, including hardware acquisition, maintenance, and energy consumption, represents a significant barrier to adoption for many manufacturing organisations (Wu et al., 2020).

Real-time constraints add another layer of complexity to the computational challenges of AI and Digital Twin applications. Many predictive maintenance applications require analysis of sensor data with latency measured in milliseconds, as delays in detecting developing faults can result in equipment damage or production disruptions (Li et al., 2021). The processing of high-frequency sensor data, such as vibration signals sampled at tens of kilohertz, presents particular challenges for real-time analysis due to the volume of data and the computational intensity of the required processing (Chen et al., 2019). The use of cloud-based processing introduces additional latency associated with data transmission, limiting the applicability of cloud computing for latency-sensitive applications (Xu et al., 2020). Edge computing architectures have emerged as a solution to these real-time challenges, enabling processing to be performed close to the data source and reducing latency (Yang et al., 2021). However, edge devices typically have limited computational capacity, constraining the complexity of AI models that can be deployed (Zhang et al., 2022). The development of model compression techniques, such as pruning, quantisation, and knowledge distillation, has shown promise for reducing the computational requirements of AI models while maintaining acceptable performance (Cheng et al., 2020).

2.6.5. Cybersecurity Concerns in Connected Systems

The increasing connectivity of manufacturing systems, which enables AI-based predictive maintenance and Digital Twin applications, also creates significant cybersecurity vulnerabilities that must be addressed to ensure safe and reliable operations. The integration of industrial equipment with IT networks and the Internet exposes previously isolated systems to cyber threats, including malware, ransomware, and targeted attacks (Nakao et al., 2020). The consequences of successful cyberattacks on manufacturing systems can be severe, including equipment damage, production disruption, intellectual property theft, and even safety incidents (Zio et al., 2021). The proliferation of IoT devices in manufacturing environments has dramatically expanded the attack surface, with many of these devices having limited security capabilities (Asghar et al., 2020). The use of legacy equipment that was not designed with modern security considerations further exacerbates these vulnerabilities, as these systems often cannot be patched or secured through conventional means (Yadav et al., 2021). The connectivity required for Digital Twins and AI-based maintenance systems creates data flows between operational technology (OT) and information technology (IT) systems that may not have been integrated in the past, introducing new attack vectors (Babu et al., 2022).

The specific cybersecurity challenges associated with AI-based predictive maintenance and Digital Twins include data integrity attacks, model poisoning, and adversarial attacks on AI systems. Data integrity attacks, where attackers manipulate sensor data to cause incorrect decisions, are particularly concerning in predictive maintenance contexts where erroneous analysis could lead to inappropriate maintenance actions or failure to detect developing faults (Patel et al., 2021). Model poisoning attacks, where attackers manipulate training data to introduce vulnerabilities or biases into AI models, can have long-lasting effects that are difficult to detect (Raman et al., 2020). Adversarial attacks, where carefully crafted inputs are designed to fool AI models, can cause predictive maintenance systems to misclassify equipment conditions, potentially leading to dangerous situations (Khalil et al., 2022). The integration of Digital Twins with enterprise systems introduces additional cybersecurity concerns related to data privacy and confidentiality, as Digital Twins contain detailed information about manufacturing processes and equipment (Fernandez et al., 2021). Addressing these cybersecurity concerns requires a comprehensive approach that includes technical controls, organisational policies, and workforce training (Khan et al., 2020). The development of cybersecurity standards and

best practices specifically for AI and Digital Twin applications in manufacturing is an emerging area that requires ongoing attention (Li et al., 2022).

2.7. Summary of Literature Review

This chapter has provided a comprehensive review of the literature pertaining to Artificial Intelligence-based Predictive Maintenance and Digital Twin development for smart manufacturing systems. The review began by examining the foundational concepts of Smart Manufacturing Systems (SMS), including the key enabling technologies of IoT, Cloud/Edge Computing, and Cyber-Physical Systems (CPS) that form the technological backbone of modern manufacturing transformation (Bianchini et al., 2024; Villalonga Jaén, 2026). The digital thread and data flow in SMS were explored, highlighting the importance of seamless data integration across the product lifecycle and the role of standards such as MTConnect, OPC UA, and MQTT in enabling bidirectional data flow between physical and digital systems (Bianchini et al., 2024). The evolution of maintenance strategies from reactive to proactive approaches was traced, demonstrating the progression from corrective and preventive maintenance to condition-based and predictive maintenance, driven by the increasing availability of data and advances in digital technologies (De Sanctis, 2023; Zio, 2024). The review also examined Condition-Based Maintenance (CBM) and Prognostics and Health Management (PHM) as sophisticated approaches that leverage real-time operational data and predictive analytics to optimise maintenance decisions (Borgi et al., 2017; Ward et al., 2024).

The literature review then explored the diverse data sources for predictive maintenance, encompassing temperature sensors, vibration sensors, sound/noise detection systems, and current/voltage monitoring devices that provide comprehensive operational and environmental status monitoring (Magadán et al., 2021; Springer, 2024). The review examined the various AI techniques applied to predictive maintenance, including machine learning approaches such as regression, classification, Support Vector Machines, and Random Forests (Susto et al., 2015; Wang et al., 2017; Widodo & Yang, 2017). Deep learning approaches including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and Autoencoders were explored, demonstrating their effectiveness for automatic feature extraction from raw sensor data and anomaly detection (Ranzato et al., 2019; Guo et al., 2016; Malhotra et al., 2016). Hybrid and ensemble models were examined, highlighting the benefits of combining multiple algorithms to achieve superior performance (Yang et al., 2020; Li et al., 2020). Deep Reinforcement Learning (DRL) for dynamic maintenance scheduling was also reviewed, demonstrating the potential of adaptive policies that optimise maintenance decisions in complex, uncertain environments (Li et al., 2021; Tian et al., 2021).

The Digital Twin concept was analysed in depth, including definitions and levels of fidelity from Digital Shadows to Digital Twins and Digital Twin Prototypes (Grieves & Vickers, 2017; Rasheed et al., 2020; Böhm et al., 2021). The key components of Digital Twins, physical asset, virtual model, data link, and services, were examined, highlighting the importance of bidirectional data flow and integration capabilities (Tao et al., 2019; Kritzinger et al., 2018; Semeraro et al., 2021). The three primary types of Digital Twins-Product DT, Production DT, and Performance DT, were explored, demonstrating the application of Digital Twins across different aspects of the product lifecycle (Díaz-Rozo et al., 2018; Cimino et al., 2019; Zheng et al., 2018). Digital Twin frameworks and architectures, including the five-dimensional model and RAMI 4.0, were examined, providing structured approaches to development and implementation (Tao et al., 2018; Weyer et al., 2016; Bader et al., 2021). The synergy between AI and Digital Twins for predictive maintenance was explored, including the use of Digital Twins as data generators for AI model training (Huang et al., 2020; Mylonas et al., 2021), AI for real-time anomaly detection (Luo et al., 2020; Mazumder et al., 2021), and AI for predictive analytics and what-if simulation (Wang et al., 2020; Singh et al., 2021).

The literature review identified several critical research gaps and challenges that limit the widespread adoption and effectiveness of AI-based predictive maintenance and Digital Twin technologies. Data quality, availability, and labelling issues persist as major barriers, with industrial environments posing significant challenges for reliable data collection and the scarcity of labelled failure data limiting the applicability of supervised learning approaches (Zheng et al., 2021; Chen et al., 2020; Li et al., 2020). Model interpretability and trust concerns, addressed through Explainable AI (XAI) approaches, remain significant challenges, as maintenance engineers are reluctant to act on recommendations they cannot understand or verify (Arrieta et al., 2020; Gunning et al., 2019; Rudin, 2019). Integration complexity and scalability challenges were identified, particularly in the context of heterogeneous legacy IT landscapes and the transition from pilot projects to full-scale deployment (Perez et al., 2021; Tiwari et al., 2022; Singh et al., 2020). Computational costs and real-time constraints present significant barriers, particularly for deep learning models and edge computing environments (Zhang et al., 2020; Chen et al., 2021; Yang et al., 2021). Cybersecurity concerns in connected systems were also

identified as critical challenges, with the increasing connectivity of manufacturing systems creating significant vulnerabilities that must be addressed (Nakao et al., 2020; Babu et al., 2022; Khalil et al., 2022).

The comprehensive review conducted in this chapter provides the foundation for developing a conceptual framework that addresses the identified gaps and challenges. The next chapter will present a theoretical framework and system architecture for AI-driven predictive maintenance within Digital Twin-enabled smart manufacturing systems, integrating insights from the literature to propose a comprehensive approach that addresses the technical, organisational, and practical challenges identified in this review.

CHAPTER THREE

Theoretical Framework and System Architecture

3.1. Introduction to the Proposed Conceptual Framework

The comprehensive literature review conducted in Chapter 2 has identified significant research gaps and challenges that limit the effective integration of Artificial Intelligence-based Predictive Maintenance with Digital Twin technology in smart manufacturing systems. These challenges include data quality and availability issues, model interpretability concerns, integration complexity, computational constraints, and cybersecurity vulnerabilities. To address these gaps, this chapter presents a novel conceptual framework that synthesises theoretical foundations from information processing theory, systems theory, and maintenance management theory to guide the development and implementation of AI-driven predictive maintenance within Digital Twin environments.

The proposed framework is grounded in Information Processing Theory (IPT), which posits that organisations and systems must develop mechanisms that can acquire, interpret, and respond to information effectively, scaling their processing capacity to match the complexity of their operating environment (Galbraith, 1974). In the context of predictive maintenance, uncertainty arising from data-level anomalies, model-level fragility, and system-level interconnectedness drives the need for sophisticated information-processing mechanisms that can transform raw sensor data into actionable maintenance insights (Jin et al., 2025). The framework applies IPT principles to maintenance systems by conceptualising intelligent fault diagnosis as a multi-layered information-processing system, enabling systematic mapping from uncertainty sources to appropriate technological and organisational countermeasures (Mailula & Saha, 2025). This theoretical grounding provides a robust foundation for designing architectures that align information-processing capacity with the complexity and uncertainty inherent in modern manufacturing environments.

The framework also draws on systems theory, recognising that manufacturing enterprises function as complex socio-technical systems comprising interconnected physical, informational, and human components (Huang et al., 2023). This systems perspective emphasises the importance of understanding interactions and dependencies between components, as mechanical breakdowns rarely occur in isolation but cascade through coupled subsystems (Xu et al., 2025). The framework therefore adopts a holistic approach that considers the entire system lifecycle, from design and development through deployment, operation, and continuous improvement. This lifecycle perspective aligns with the principles of Maintenance 5.0, which emphasises sustainability, resilience, and human-centricity alongside technological advancement (Chen et al., 2025).

3.2. System Architecture for AI-Driven PdM with Digital Twin

3.2.1. The Physical Layer (Sensors and Equipment)

The physical layer constitutes the foundation of the proposed architecture, encompassing all tangible assets, equipment, sensors, and actuators that comprise the manufacturing system. This layer includes production machinery, robots, conveyors, measurement devices, and the sensors that monitor their operational status and environmental conditions. Typical sensors employed in this layer include vibration sensors for mechanical condition monitoring, temperature sensors for thermal analysis, current and voltage sensors for electrical signature analysis, pressure sensors for fluid systems, and acoustic sensors for sound-based fault detection (Magadán et al., 2021). The selection and placement of sensors must be guided by a thorough understanding of equipment failure modes and the physical phenomena that precede failure, ensuring that critical degradation mechanisms are captured by appropriate monitoring technologies.

The physical layer is characterised by its heterogeneity, encompassing equipment from different vendors, with varying communication protocols, data formats, and operational characteristics. This heterogeneity presents significant challenges for data integration that must be addressed through standardisation and abstraction mechanisms (Bianchini et al., 2024). The physical layer also imposes practical constraints on data collection, including limitations on sensor placement, power availability, and communication bandwidth that must be considered in the overall system design. The reliability of data from the physical layer is fundamental to the effectiveness of the entire system, as the quality of predictions and decisions depends directly on the quality and representativeness of the underlying sensor data (Zheng et al., 2021).

3.2.2. The Data Acquisition and Communication Layer (Edge to Cloud)

The data acquisition and communication layer bridges the physical and digital domains, implementing the mechanisms for collecting, transporting, and initial processing of sensor data from the physical layer. This layer encompasses the Industrial Internet of Things (IIoT) infrastructure, including sensors, gateways, edge computing devices, and communication networks that

enable the flow of data from the factory floor to the digital twin and analytical platforms. Standard communication protocols such as OPC UA and MQTT are employed to ensure interoperability and reliable data transmission between heterogeneous devices and systems (Bianchini et al., 2024). Edge computing devices perform initial data preprocessing, including filtering, aggregation, and normalisation, reducing the volume of data that must be transmitted to central systems and enabling low-latency responses for time-critical applications (Villalonga Jaén, 2026).

The data acquisition and communication layer implements the "edge-to-cloud continuum" principle, distributing processing capabilities across the network to balance latency, bandwidth, and computational requirements (Bianchini et al., 2024). Edge devices handle time-sensitive processing tasks, such as real-time anomaly detection and immediate alert generation, while cloud or on-premise servers handle more computationally intensive tasks such as model training, complex simulations, and long-term data analysis. This distributed architecture ensures that the system can meet the real-time constraints of predictive maintenance applications while leveraging the scalability and processing power of cloud computing for strategic decision-making (Yang et al., 2021). The communication infrastructure must be designed with redundancy and resilience to ensure continuous data flow even in the event of network disruptions or component failures.

3.2.3. The Data Management and Processing Layer (Data Lake/Historian)

The data management and processing layer provides the infrastructure for storing, organising, and managing the vast volumes of data generated by manufacturing operations. This layer typically implements a data lake or data historian architecture that can accommodate structured and unstructured data from diverse sources, including time-series sensor data, maintenance records, production logs, and quality measurements. The data lakehouse architecture, which combines the flexibility of data lakes with the performance and governance capabilities of data warehouses, has emerged as a preferred approach for manufacturing data management (Bianchini et al., 2024). This architecture supports the storage of raw data for future analysis, as well as processed and aggregated data for operational applications.

The data management layer also implements data governance practices that ensure data quality, consistency, and security across the system. These practices include data validation and cleaning procedures that address issues such as missing values, outliers, and sensor drift that are common in industrial environments (Zheng et al., 2021). Metadata management capabilities enable the tracking of data provenance, lineage, and context, facilitating the interpretation of data by both human analysts and machine learning algorithms. The data management layer also provides interfaces for data access and querying, enabling analytical applications and digital twins to retrieve the data they require for their respective functions (Bianchini et al., 2024). The performance and scalability of the data management layer are critical to the overall system effectiveness, as delays or failures in data access can compromise real-time monitoring and decision-making capabilities.

3.2.4. The Digital Twin Layer (Virtual Replication and Simulation)

The digital twin layer is the core of the proposed architecture, providing a dynamic virtual representation of physical assets that synchronises with real-time sensor data and supports simulation, analysis, and decision-making. This layer encompasses multiple digital twin instances, each corresponding to a physical asset or system, maintained throughout its lifecycle (Tao et al., 2018). The digital twin layer implements the five-dimensional framework proposed by Tao et al., encompassing the physical entity (PE), virtual entity (VE), services (Ss), connection (CN), and data (DD) dimensions that together enable comprehensive modelling and analysis capabilities (Tao et al., 2019). The virtual entity includes geometric models for visualisation, physics-based models for simulation, and data-driven models for prediction, integrated into a coherent representation of the physical asset.

The digital twin layer enables real-time synchronisation between the physical and virtual realms, ensuring that the digital representation accurately reflects the current state of the physical asset (Kritzinger et al., 2018). This synchronisation is achieved through the continuous ingestion of sensor data from the data acquisition layer, with the digital twin updating its state in response to changes in the physical world. The digital twin layer also supports what-if simulation, enabling the exploration of alternative scenarios and the evaluation of different maintenance strategies before they are applied to the physical system (Tao et al., 2019). The Asset Administration Shell (AAS) framework provides a standardised approach to implementing digital twins, establishing common data structures and interfaces that enable interoperability across different vendors and systems (Bader et al., 2021). The digital twin layer thereby serves as a bridge between the physical and analytical domains, enabling the application of advanced analytics to complex industrial problems.

3.2.5. The AI/Analytics Layer (Model Development, Training, and Inference)

The AI and analytics layer provides the intelligence that drives predictive maintenance decision-making,

implementing the machine learning and deep learning models that analyse sensor data to detect anomalies, predict failures, and recommend maintenance actions. This layer encompasses the entire machine learning lifecycle, from data preparation and feature engineering through model development, training, validation, and deployment. The MLOps (Machine Learning Operations) practices are employed to manage this lifecycle systematically, ensuring that models are developed, tested, and deployed in a repeatable and reliable manner (Bianchini et al., 2024). Feature stores provide centralised repositories of derived features that can be reused across different models, improving consistency and reducing development effort (Bianchini et al., 2024).

The AI and analytics layer implements a range of analytical techniques appropriate to different maintenance tasks and data characteristics. Supervised learning techniques, including regression and classification, are employed when labelled historical data is available to train models for failure prediction and fault classification (Susto et al., 2015). Unsupervised learning techniques, including autoencoders and clustering, are employed for anomaly detection when labelled failure data is scarce (Sakurada & Yairi, 2016). Deep learning techniques, including Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, are employed for automatic feature extraction from raw sensor data and for modelling complex temporal patterns (Guo et al., 2016; Malhotra et al., 2016). Ensemble methods combine multiple models to improve prediction accuracy and robustness (Yang et al., 2020). The AI and analytics layer also includes model monitoring and retraining capabilities, addressing the challenge of model drift as equipment and operating conditions evolve over time (Zhao et al., 2022).

3.2.6. The Application/Visualization Layer (Dashboards, Alerts, Decision Support)

The application and visualisation layer translates the outputs of the AI and digital twin layers into actionable insights for human decision-makers, providing the interfaces through which operators, engineers, and managers interact with the system. This layer encompasses dashboards for real-time monitoring of equipment health and performance, alert systems for notifying personnel of developing issues, and decision support tools for evaluating alternative maintenance strategies. Modern visualisation approaches leverage game engine technologies to provide immersive, intuitive representations of the manufacturing system, enabling operators to understand the location and nature of detected issues (Bianchini et al., 2024). Augmented Reality (AR) interfaces can overlay maintenance information on the physical equipment, guiding operators through inspection and repair procedures.

The application and visualisation layer implements Explainable AI (XAI) techniques to enhance the interpretability and trustworthiness of AI predictions. Feature importance visualisations, such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations), provide insights into which factors most influenced a particular prediction, helping operators understand the reasoning behind model outputs (Lundberg & Lee, 2017; Ribeiro et al., 2016). Visualisation of model confidence and uncertainty enables risk-informed decision-making, allowing operators to weigh the reliability of predictions against the consequences of different actions (Samek et al., 2021). The application and visualisation layer also supports collaborative decision-making, enabling maintenance, production, and management teams to share information and coordinate their actions based on a common understanding of the system state.

3.3. Digital Twin Design and Modeling Methodologies

3.3.1. Geometric and Physical Modeling (CAD, FEA)

Geometric and physical modelling provides the structural foundation for the digital twin, capturing the geometry, material properties, and physical characteristics of equipment and systems. Computer-Aided Design (CAD) models are employed to create accurate geometric representations of equipment, providing visual fidelity for the digital twin and enabling the visualisation of asset location and spatial relationships (Schleich et al., 2017). These models may range from simple wireframe representations to detailed solid models that capture complex geometries and assemblies. The level of geometric detail required depends on the intended application, with visualisation and training applications requiring less detail than applications requiring precise physical simulation.

Finite Element Analysis (FEA) extends geometric modelling by capturing the physical behaviour of equipment under various loading conditions. FEA models solve partial differential equations that describe physical phenomena such as stress, strain, vibration, and heat transfer, providing detailed insights into equipment behaviour that are not available from sensor data alone (Zheng et al., 2019). These physics-based models are particularly valuable for understanding degradation mechanisms and predicting failure modes, as they can simulate the effects of different operating conditions and failure scenarios (Raissi et al., 2019). Physics-informed neural networks (PINNs) offer a promising approach for integrating physics-based and data-driven modelling, embedding physical constraints into the learning process to improve prediction accuracy and generalisation (Raissi et al., 2019). The combination of CAD and FEA models with sensor data enables the creation of high-fidelity digital

twins that accurately represent both the geometry and behaviour of physical assets.

3.3.2. Behavioral and Simulation Modeling (Discrete-Event, Agent-Based, System Dynamics)

Behavioural and simulation modelling captures the dynamic behaviour of manufacturing systems, enabling the simulation of system operation, the evaluation of different scenarios, and the optimisation of system performance. Discrete-Event Simulation (DES) models represent the operation of systems as a sequence of discrete events, such as part arrivals, machine operations, and equipment failures, enabling the simulation of complex production systems and the evaluation of alternative configurations and operating policies (Banks et al., 2010). DES models are particularly valuable for evaluating the impact of maintenance policies on system throughput and availability, as they can capture the interactions between maintenance activities and production operations (Mourtzis et al., 2019).

Agent-Based Modelling (ABM) provides an alternative approach to behavioural simulation, representing the system as a collection of autonomous agents that interact according to defined rules. This approach is particularly suited to modelling distributed, adaptive, and emergent behaviours that are difficult to capture with DES models (Bonabeau, 2002). In the context of predictive maintenance, ABM can represent the interactions between different equipment components, maintenance resources, and production demands, capturing the emergent behaviours that arise from these interactions (Macal & North, 2010). System Dynamics (SD) modelling provides a high-level approach to behavioural simulation, representing the system as a set of stocks, flows, and feedback loops that capture the dynamic relationships between key variables (Sterman, 2000). SD models are particularly valuable for understanding long-term system behaviour and evaluating the effects of strategic decisions.

3.3.3. Data-Driven Modeling and the Role of AI in DT Creation

Data-driven modelling leverages sensor data and machine learning techniques to create digital twin models that capture the behaviour of equipment and systems without explicit representation of underlying physical mechanisms. This approach is particularly valuable when physical models are unavailable, too complex, or too computationally intensive for real-time applications. Data-driven models may employ a range of techniques, including regression models for predicting key performance indicators, neural networks for capturing complex non-linear relationships, and statistical models for characterising uncertainty (Zio, 2024). The data-driven approach enables the continuous updating of models as new data becomes available, maintaining accuracy even as equipment degrades and operating conditions evolve (Lee et al., 2015).

The role of AI in digital twin creation extends beyond model building to encompass the entire lifecycle of digital twin development and maintenance. Machine learning techniques are employed for model calibration, using sensor data to adjust physical model parameters to match observed behaviour (Huang et al., 2020). Generative AI techniques are employed for synthetic data generation, creating training data for AI models when real failure data is scarce (Mylonas et al., 2021). AutoML techniques automate the process of model selection and hyperparameter optimisation, accelerating the development of accurate predictive models (Wang et al., 2021). The integration of AI techniques throughout the digital twin lifecycle is essential for creating and maintaining accurate, up-to-date digital representations of complex manufacturing systems (Tao et al., 2019).

3.4. AI Model Selection and Development Methodology

3.4.1. Data Preprocessing and Feature Engineering for Multi-Sensor Data

The effectiveness of AI models for predictive maintenance depends critically on the quality of the data they are trained on and the features they are provided. Data preprocessing addresses issues such as missing values, outliers, noise, and sensor drift that are common in industrial sensor data (Zheng et al., 2021). Missing values may be handled through interpolation, forward filling, or model-based imputation depending on the pattern and extent of missingness (Chen et al., 2020). Outliers may be detected and removed or corrected using statistical techniques such as the Interquartile Range (IQR) method or Z-score analysis. Noise reduction may be achieved through filtering techniques such as moving averages or more sophisticated approaches such as wavelet denoising.

Feature engineering transforms raw sensor data into features that are more informative for predictive modelling. Domain knowledge is essential for identifying features that capture the physical mechanisms of equipment degradation and failure (Lee et al., 2015). Common features for predictive maintenance include statistical features such as mean, variance, skewness, and kurtosis; frequency-domain features such as power spectral density and harmonic content; and time-frequency features such as wavelet coefficients (Jardine et al., 2006). Feature extraction from multi-sensor data requires careful alignment of data from different sources, accounting for differences in sampling rates, time offsets, and measurement scales (Monostori et al., 2016). Feature selection techniques such as Principal

Component Analysis (PCA) and feature importance analysis may be employed to reduce the dimensionality of the feature space and improve model performance (Li et al., 2020).

3.4.2. Selection Criteria for Algorithms

The selection of appropriate AI algorithms for predictive maintenance applications requires systematic consideration of the problem characteristics, data availability, and operational constraints. Table 1 summarises the key factors that should guide algorithm selection, including the nature of the prediction task (classification, regression, anomaly detection), the volume and type of data available (labelled, unlabelled, time-series, image), and the requirements for model interpretability and computational efficiency (Gokalp et al., 2020).

Table 1: Algorithm Selection Criteria for Predictive Maintenance Applications

Criterion	Description	Implications for Algorithm Selection
Prediction Task	Classification, Regression, Anomaly Detection	Classification tasks require algorithms such as SVM, Random Forest; Regression tasks require algorithms such as LSTM; Anomaly detection requires algorithms such as Autoencoders
Data Availability	Labelled vs. Unlabelled Data	Supervised learning requires labelled data; Semi-supervised or unsupervised learning can be used when labelled data is scarce
Data Type	Time-series, Image, Multivariate	Time-series data requires algorithms such as LSTM; Image data requires algorithms such as CNN
Interpretability Requirements	High vs. Low	Interpretable models such as Decision Trees or Rule-based systems may be preferred for safety-critical applications
Computational Constraints	Training/Inference time, Memory	Edge deployment may require lightweight models such as Random Forest or quantised neural networks
Accuracy Requirements	Required accuracy level for safe operation	Complex models such as Ensemble Methods may be required for high-accuracy applications

The nature of the prediction task is the primary determinant of algorithm selection. Classification tasks, such as distinguishing between normal and faulty operation, may be addressed with algorithms such as Support Vector Machines (SVM), Random Forests, or Convolutional Neural Networks (CNNs) (Widodo & Yang, 2017; Zhao et al., 2018). Regression tasks, such as estimating remaining useful life, may be addressed with algorithms such as Long Short-Term Memory (LSTM) networks or Gaussian Process Regression (Ren et al., 2018; Wang et al., 2017). Anomaly detection tasks, where the goal is to identify unusual patterns that may indicate developing faults, may be addressed with algorithms such as Autoencoders or Isolation Forests (Sakurada & Yairi, 2016; Liu et al., 2008).

Interpretability requirements are particularly important in safety-critical applications, where operators must understand the reasoning behind model predictions to make informed decisions (Arrieta et al., 2020). When interpretability is required, simpler models such as Decision Trees, Random Forests, or Logistic Regression may be preferred over more complex "black box" models (Rudin, 2019). However, the trade-off between interpretability and accuracy must be carefully considered, as more interpretable models may have lower predictive performance (Caruana et al., 2015). Computational constraints may also influence algorithm selection, particularly for edge-based applications where computational resources are limited (Yang et al., 2021). In such cases, lightweight models such as Random Forests or quantised neural networks may be preferred over more computationally intensive deep learning models.

3.4.3. Model Training, Validation, and Hyperparameter Optimization

The model training and validation process is critical to ensuring the reliability and generalisability of AI models for predictive maintenance. The process begins with the splitting of available data into training, validation, and test sets, ensuring that the test set is reserved for final evaluation and not used for model development (Susto et al., 2015). Cross-validation techniques, such as k-fold cross-validation, are employed to provide robust estimates of model performance and to guard against overfitting (Hastie et al., 2009). Time-series data requires special handling during train-test splitting to avoid data leakage, with training and test sets separated by time rather than sampled randomly (Zhang et al., 2017).

Hyperparameter optimisation is employed to find the optimal configuration of model parameters that maximise performance on the validation set. Grid search and random search provide systematic approaches to exploring the hyperparameter space, while more sophisticated approaches such as Bayesian optimisation and automated machine learning (AutoML) can reduce the computational cost of hyperparameter search (Bergstra & Bengio, 2012; Snoek et al., 2012). The choice of performance metrics is also important, with metrics such as accuracy, precision, recall, F1-score, and Area Under the Receiver Operating Characteristic Curve (AUC-ROC) providing different perspectives on model performance (Sokolova & Lapalme, 2009). In predictive maintenance applications, where the cost of false negatives (missed failures) may be much higher than the cost of false positives (false alarms), metrics that emphasise recall may be more appropriate than overall accuracy (Gokalp et al., 2020).

3.4.4. Model Deployment and Continuous Learning (MLOps)

The deployment and ongoing maintenance of AI models in production environments is managed through MLOps practices that extend DevOps principles to machine learning applications (Bianchini et al., 2024). Model deployment involves the packaging of trained models into deployable artefacts, the validation of model performance in a staging environment, and the controlled rollout to production (Gao et al., 2021). Continuous Integration/Continuous Deployment (CI/CD) pipelines automate the process of model development, testing, and deployment, enabling rapid iteration while maintaining quality and reliability (Bianchini et al., 2024). Model monitoring tracks model performance in production, detecting degradation due to data drift or concept drift that may require model retraining (Zhao et al., 2022).

Continuous learning addresses the challenge of model drift by enabling models to be updated as new data becomes available. Batch retraining approaches periodically retrain models using historical data, while online learning approaches update models incrementally as each new data point arrives (Gao et al., 2021). The choice between batch and online learning depends on the rate of change in the underlying data distribution and the computational resources available for retraining (Liu et al., 2021). Model versioning and experiment tracking ensure that the development and deployment process is reproducible and auditable, supporting the traceability requirements of regulated industries (Bianchini et al., 2024). The integration of MLOps practices with digital twin and predictive maintenance systems is essential for maintaining the accuracy and reliability of AI models over time.

3.5. Integrating AI with the Digital Twin: A Bidirectional Feedback Loop

3.5.1. Predictive Insights Feeding the DT for Simulation

The integration of AI predictions with digital twin simulations enables a bidirectional feedback loop that enhances the capabilities of both systems. Predictive insights from AI models feed into the digital twin, providing information about future equipment health and performance that can be used to inform simulations of future scenarios (Tao et al., 2018). This integration enables predictive simulation, where the digital twin uses AI predictions to forecast the future state of the system and evaluate the consequences of different actions (Wang et al., 2020). For example, remaining useful life predictions can be used to simulate the effects of different maintenance timing decisions, enabling the optimisation of maintenance schedules based on predicted future conditions (Singh et al., 2021).

The integration of predictive insights into the digital twin also enables uncertainty quantification and scenario analysis. AI predictions are inherently uncertain, due to the limitations of training data, the stochastic nature of degradation processes, and the complexity of manufacturing environments (Rasheed et al., 2020). By incorporating uncertainty estimates into the digital twin, operators can explore the range of possible future outcomes and evaluate the risk associated with different decisions (Bottani et al., 2022). Scenario analysis enables the exploration of alternative possible futures, helping organisations prepare for a range of outcomes and build resilience against unexpected events (Hao et al., 2021). The integration of predictive insights into the digital twin thereby transforms the digital twin from a descriptive tool into a predictive and prescriptive tool that supports proactive maintenance decision-making.

3.5.2. "What-If" Simulation Results for Maintenance Optimization

"What-if" simulation within the digital twin enables the evaluation of alternative maintenance strategies, supporting the optimisation of maintenance decisions based on predicted outcomes (Olivares et al., 2022). The simulation process begins with the definition of alternative maintenance actions or policies, which may include different timing of interventions, different maintenance scopes (repair vs. replace), or different resource allocations (Singh et al., 2021). The digital twin simulates the effects of each alternative, predicting the future health and performance of the equipment under each scenario (Hao et al., 2021). The results of these simulations are compared to identify the optimal maintenance strategy, balancing the

costs of maintenance against the risks of failure and the value of continued operation (Chen et al., 2022).

The what-if simulation capability also supports the optimisation of maintenance resource allocation, enabling organisations to make the most effective use of limited maintenance resources. By simulating the effects of different resource allocation strategies, organisations can identify the most critical assets to prioritise, the optimal timing of interventions, and the most effective combination of maintenance activities (Sun et al., 2021). The iterative nature of what-if simulation enables continuous improvement, with the outcomes of actual maintenance decisions compared against simulation predictions to refine models and improve future decision-making (Xu et al., 2021). The visualisation capabilities of the digital twin enable stakeholders to understand the implications of different maintenance strategies intuitively, supporting collaborative decision-making among maintenance, production, and management teams (Lee et al., 2021). The what-if simulation capability thereby provides a powerful tool for moving from reactive to proactive maintenance management.

3.6. Considerations for Explainability and Trust (XAI)

3.6.1. Techniques for Interpretability (e.g., SHAP, LIME, Attention Mechanisms)

Explainable Artificial Intelligence (XAI) techniques are essential for building trust in AI-based predictive maintenance systems and enabling informed decision-making. Model-agnostic techniques, such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations), provide insights into model predictions for any type of model, enabling comparison and verification across different algorithms (Lundberg & Lee, 2017; Ribeiro et al., 2016). SHAP provides a unified framework for feature importance analysis, assigning each feature a contribution value based on game-theoretic principles that ensure consistency and fairness (Lundberg & Lee, 2017). LIME provides local explanations for individual predictions, fitting interpretable models to the local neighbourhood of a prediction to explain what features were most influential in that specific case (Ribeiro et al., 2016).

Model-specific interpretability techniques are available for particular classes of models. For decision trees and random forests, feature importance scores and visualisation of decision paths provide insights into model behaviour (Zhao et al., 2018). For neural networks, techniques such as saliency maps, layer-wise relevance propagation (LRP), and attention visualisation provide insights into which input features or internal representations are most influential for predictions (Zeiler & Fergus, 2014; Bach et al., 2015). For LSTM networks, attention mechanisms can be visualised to show which time steps the model is focusing on for its predictions, helping identify the temporal patterns associated with different fault conditions (Choi et al., 2018). The choice of interpretability technique depends on the model type, the nature of the data, and the needs of the stakeholders who will be using the explanations.

3.6.2. Role of the DT in Visualizing the Reasoning of the AI

The digital twin provides a powerful platform for visualising the reasoning of AI models, making complex predictions more understandable and actionable for human operators (Mazumder et al., 2021). By overlaying AI predictions and explanations on the virtual representation of the physical system, the digital twin enables operators to see not only what the AI predicts, but also why it makes those predictions (Park et al., 2021). For example, anomaly detection results can be visualised on the digital twin, highlighting the specific components that are exhibiting anomalous behaviour and the magnitude of the deviation from normal (Kim et al., 2022). The visualisation of feature importance on the digital twin can show which sensor measurements are most influential for a particular prediction, helping operators understand the physical phenomena driving the AI's assessment (Perez et al., 2022).

The digital twin also supports the explanation of AI predictions through simulation and what-if analysis. By showing how predicted outcomes change under different scenarios, the digital twin can help operators understand the conditions under which predictions are likely to be accurate and the factors that influence different outcomes (Hao et al., 2021). The integration of uncertainty visualisation into the digital twin enables operators to understand the confidence of AI predictions, supporting risk-informed decision-making (Bottani et al., 2022). The digital twin thereby serves as a bridge between the abstract world of AI predictions and the concrete world of physical equipment, making AI more accessible and trustworthy for industrial practitioners.

CHAPTER FOUR

Analysis And Synthesis of Implementation Scenarios

4.1. Introduction

The theoretical framework and system architecture proposed in Chapter 3 provide a comprehensive foundation for understanding how Artificial Intelligence-based Predictive Maintenance and Digital Twin technologies can be integrated within smart manufacturing systems. However, the translation of theoretical constructs into practical implementation requires careful consideration of the specific contexts, challenges, and opportunities that characterise different manufacturing domains. This chapter addresses this need through the analysis and synthesis of three distinct implementation scenarios that represent key application areas for AI-driven predictive maintenance and Digital Twin technologies.

The scenarios examined in this chapter have been selected to represent different levels of complexity, different types of equipment, and different operational contexts within manufacturing. Scenario 1 focuses on predictive maintenance for critical rotating equipment, representing the application of these technologies to high-value, mission-critical assets where unplanned failures have significant consequences for safety, production, and cost (Phutane & Pawar, 2025). Scenario 2 addresses production line throughput optimisation and predictive quality, representing the application of these technologies at the system level where interactions between multiple pieces of equipment and processes must be considered (Yadin & Cheng, 2025). Scenario 3 examines collaborative robots and human-robot interaction, representing the emerging application of these technologies to systems where human workers and automated systems must work together safely and effectively (Mazumder et al., 2023).

The analysis of each scenario follows a consistent structure that examines the Digital Twin development approach, the AI applications employed, and the simulation and what-if analysis capabilities enabled by the integration of these technologies. This structure enables systematic comparison across scenarios and the identification of common patterns, architectures, and success factors that can inform implementation decisions across different manufacturing contexts. The chapter concludes with a cross-case analysis that synthesises findings from the three scenarios, identifying common patterns in data requirements, Digital Twin fidelity levels, AI algorithm selection, and integration approaches. A generic implementation roadmap is proposed that translates these findings into actionable guidance for manufacturing organisations considering the adoption of these technologies.

4.2. Scenario 1: Predictive Maintenance for Critical Rotating Equipment (e.g., Turbines, Compressors)

4.2.1. DT Development: Physics-Based Model Augmented with Sensor Data

Critical rotating equipment, including gas turbines, compressors, pumps, and motors, represent some of the most valuable and mission-critical assets in manufacturing and process industries. The development of Digital Twins for such equipment must address the unique characteristics of rotating machinery, including complex multi-physics interactions between electromagnetic, mechanical, and thermal domains (Villalonga Jaén, 2026). The Digital Twin development approach for this scenario typically begins with the creation of physics-based models that capture the fundamental behaviour of the equipment under various operating conditions. These models are developed using Finite Element Analysis (FEA) for structural analysis, Computational Fluid Dynamics (CFD) for fluid flow analysis, and multi-body dynamics for mechanical system analysis (Aivaliotis et al., 2021).

The physics-based models are then augmented with sensor data to create hybrid Digital Twins that combine the strengths of physics-based and data-driven approaches. The sensor data typically includes vibration measurements from accelerometers, temperature measurements from thermocouples, pressure measurements from transducers, and current and voltage measurements from electrical monitoring systems (Hosamo et al., 2022). The sensor data serves two critical functions: it provides real-time updates that synchronise the Digital Twin with the physical asset, and it enables the calibration of physics-based model parameters to match observed behaviour (Rasheed et al., 2020). This calibration process, sometimes referred to as model updating or data assimilation, is essential for maintaining the accuracy of the Digital Twin as equipment degrades and operating conditions evolve (Aivaliotis et al., 2021).

The architecture of a rotating equipment Digital Twin typically follows a layered structure that separates different functions and concerns. The physical layer encompasses the equipment itself, including all sensors and actuators that enable monitoring and control. The communication layer implements the data acquisition and transmission infrastructure, typically using industrial protocols such as OPC UA or MQTT for sensor data collection (Bianchini et al., 2024). The data layer provides storage and management of sensor data, maintenance records, and operational logs. The modelling layer encompasses the physics-based and data-driven models that constitute the digital representation

of the equipment. The application layer provides the visualisation, analysis, and decision support capabilities that enable maintenance planning and execution (Phutane & Pawar, 2025).

The fidelity of the Digital Twin for rotating equipment must be carefully calibrated to the requirements of the application. Higher-fidelity models that capture detailed physical behaviour provide more accurate predictions and enable more sophisticated what-if analysis, but they also require more computational resources and may not be suitable for real-time applications (Lee et al., 2019). Lower-fidelity models, such as reduced-order models or surrogate models, can be used for real-time monitoring and anomaly detection, while higher-fidelity models can be reserved for detailed analysis and planning (Schleich et al., 2017). The Digital Twin architecture must therefore support the simultaneous deployment of multiple models at different levels of fidelity, enabling the selection of the appropriate model for each task (Kritzinger et al., 2018).

4.2.2. AI Application: Vibration Analysis Using CNNs for Fault Classification

Vibration analysis is the most widely used technique for condition monitoring of rotating equipment, as vibration signatures provide rich information about the mechanical condition of bearings, gears, shafts, and other rotating components (Jardine et al., 2006). The application of AI to vibration analysis has significantly enhanced the capabilities of condition monitoring systems, enabling the automatic detection and classification of faults that might be missed by traditional threshold-based approaches. Convolutional Neural Networks (CNNs) have emerged as particularly effective for vibration analysis, as they can learn hierarchical features from vibration signals that correspond to different fault types and severity levels (Guo et al., 2016; Zhao et al., 2019). CNNs applied to vibration data typically operate on spectrograms or other time-frequency representations that preserve both the temporal and frequency structure of the signal (Wang et al., 2018).

The application of CNNs to vibration analysis for rotating equipment typically follows a systematic process that includes data acquisition, preprocessing, feature extraction, model training, and deployment. Data acquisition involves collecting vibration signals from accelerometers mounted on the equipment, typically at sampling rates of 10-20 kHz to capture the high-frequency components associated with bearing and gear faults (Chen et al., 2017). Preprocessing includes filtering to remove noise and interference, segmentation into time windows for analysis, and transformation into frequency or time-frequency domains for feature extraction (Janssens et al., 2016). Feature extraction may involve the calculation of statistical features, the generation of spectrograms using Short-Time Fourier Transform (STFT), or the application of wavelet transforms to capture time-frequency characteristics (Chen et al., 2017).

The CNN architecture for vibration classification typically consists of multiple convolutional layers that learn increasingly abstract features from the input data, followed by fully connected layers that perform classification (Guo et al., 2016). The convolutional layers apply filters to the input data, detecting patterns such as frequency peaks, harmonics, and sidebands that are characteristic of different fault types. The learned features are then fed into the classification layers, which output probabilities for each fault type or health condition (Zhao et al., 2019). Training of the CNN requires labelled data that includes examples of each fault type, which may be obtained from historical data, controlled experiments, or synthetic data generated by Digital Twins (Huang et al., 2020). The trained CNN can then be deployed for real-time vibration analysis, providing continuous monitoring and early warning of developing faults (Phutane & Pawar, 2025).

The performance of CNN-based vibration analysis for rotating equipment has been demonstrated in numerous studies, with accuracy rates typically exceeding 95% for fault classification tasks (Guo et al., 2016; Zhao et al., 2019). CNNs have been shown to be particularly effective for bearing fault classification, gear fault detection, and motor fault diagnosis (Chen et al., 2017; Janssens et al., 2016). The ability of CNNs to learn features automatically from raw data eliminates the need for manual feature engineering, reducing the expertise required to deploy effective condition monitoring systems (Wang et al., 2018). However, the effectiveness of CNN-based approaches depends on the availability of sufficient labelled training data, which may be limited for rare fault types or novel equipment (Li et al., 2020). The integration of CNNs with Digital Twins addresses this challenge by enabling the generation of synthetic training data that augments limited real-world data (Huang et al., 2020).

4.2.3. Simulation of Failure Modes

The integration of AI predictions with Digital Twin simulation capabilities enables the simulation of failure modes and the evaluation of maintenance strategies for rotating equipment. The simulation of failure modes involves the creation of scenarios that represent different types and severities of faults, enabling the analysis of their effects on equipment performance and the development of appropriate maintenance responses (Bottani et al., 2022). The Digital Twin simulates the behaviour of the equipment under fault conditions, generating predictions of vibration signatures, temperature profiles, and other indicators that can be compared with actual sensor data for fault diagnosis (Phutane & Pawar, 2025). The simulation of failure modes

is particularly valuable for equipment where historical failure data is limited, as it enables the generation of synthetic data that can be used for AI model training and validation (Huang et al., 2020).

The simulation of failure modes for rotating equipment typically involves the modification of physics-based model parameters to represent different fault conditions. For example, bearing faults may be simulated by introducing changes to bearing stiffness or damping coefficients, gear faults may be simulated by modifying gear mesh stiffness, and rotor faults may be simulated by introducing imbalances or misalignments (Aivaliotis et al., 2021). The modified models are then used to simulate the response of the equipment under various operating conditions, generating predictions of vibration, temperature, and other performance indicators (Rasheed et al., 2020). The simulated data provides a rich source of labelled training data for AI models, enabling the development of robust fault detection and classification algorithms (Huang et al., 2020).

The simulation of failure modes also enables the evaluation of alternative maintenance strategies through what-if analysis. The Digital Twin can simulate the effects of different maintenance interventions, such as bearing replacement, shaft alignment, or lubrication changes, on equipment performance and remaining useful life (Chen et al., 2022). The simulation results can be used to compare the costs and benefits of different maintenance strategies, identifying the optimal timing and scope of interventions (Singh et al., 2021). The what-if analysis capability supports proactive maintenance planning, enabling organisations to move from reactive to predictive maintenance approaches (Hao et al., 2021). The iterative nature of what-if simulation enables continuous improvement, as the outcomes of actual maintenance decisions are compared against simulation predictions to refine models and improve future decision-making (Xu et al., 2021).

The integration of failure mode simulation with AI predictions also supports the development of explainable AI (XAI) capabilities for rotating equipment maintenance. By simulating the effects of different fault conditions and visualising the resulting changes in equipment behaviour, the Digital Twin can help operators understand the reasoning behind AI predictions (Mazumder et al., 2021). For example, the Digital Twin can show how a particular vibration pattern is associated with a specific bearing fault, enabling operators to verify the AI's assessment and understand the physical mechanisms underlying the predicted failure (Kim et al., 2022). This explanatory capability is essential for building trust in AI-based maintenance systems and enabling informed decision-making (Arrieta et al., 2020).

4.3. Scenario 2: Production Line Throughput Optimization and Predictive Quality

4.3.1. DT Development: Discrete-Event Simulation Model of the Production Line

Production line optimisation and predictive quality represent a more complex application of AI and Digital Twin technologies than equipment-level predictive maintenance, as they require modelling of the interactions between multiple pieces of equipment, material flows, and human operators. The Digital Twin for a production line typically employs Discrete-Event Simulation (DES) as the primary modelling approach, as DES captures the discrete, event-driven nature of production systems where entities (parts, products) flow through a network of operations (Banks et al., 2010). The DES model represents the production line as a set of resources (machines, workers), buffers (storage areas), and flows (material, information), with events such as part arrivals, machine operations, and equipment failures triggering changes in system state (Mourtzis et al., 2019).

The development of a production line Digital Twin typically begins with the construction of a baseline DES model that represents the current production system and its operating characteristics. The baseline model is calibrated using historical production data, including cycle times, downtime statistics, and quality metrics, to ensure that it accurately reflects the behaviour of the actual production line (Banks et al., 2010). The model is validated through comparison with actual production data, with model parameters adjusted until the model outputs match observed performance within acceptable tolerances (Mourtzis et al., 2019). The validated baseline model serves as the foundation for the Digital Twin, providing a platform for simulation, analysis, and optimisation of the production line (Bottani et al., 2022).

The production line Digital Twin is then augmented with real-time data from the physical production system to enable dynamic synchronisation and what-if analysis. Real-time data from sensors, Manufacturing Execution Systems (MES), and other information systems is integrated into the Digital Twin, enabling the model to reflect the current state of the production system (Bianchini et al., 2024). This integration enables the Digital Twin to support real-time decision-making, such as dynamic scheduling of production orders, reallocation of resources in response to disruptions, and optimisation of material flow (Tao et al., 2018). The integration of real-time data also enables the continuous calibration of model parameters, ensuring that

the Digital Twin remains accurate as equipment degrades and operating conditions evolve (Aivaliotis et al., 2021).

The production line Digital Twin also encompasses the integration of predictive quality models that relate process parameters to product quality characteristics. The integration of predictive quality models with DES enables the simulation of quality outcomes under different operating scenarios, supporting the optimisation of process parameters to achieve quality targets (Bottani et al., 2022). The predictive quality models may be developed using machine learning techniques, including regression models for predicting quality characteristics, classification models for predicting defect occurrence, and deep learning models for complex quality prediction tasks (Mourtzis et al., 2019). The integration of predictive quality with DES supports proactive quality management, enabling organisations to identify and mitigate quality issues before they result in defects or rework (Yadin & Cheng, 2025).

4.3.2. AI Application: Reinforcement Learning for Dynamic Scheduling and LSTM for Quality Prediction

The complexity of production line operations, with multiple interacting processes, variable demand patterns, and uncertain equipment behaviour, makes production scheduling a challenging optimisation problem. Reinforcement Learning (RL) has emerged as a promising approach for dynamic scheduling in production environments, as it enables the learning of adaptive policies that respond to changing conditions (Li et al., 2021). RL agents learn scheduling policies through interaction with the production environment, receiving rewards for good decisions and penalties for poor decisions, gradually improving performance over time (Chen et al., 2021). The RL agent observes the current state of the production system, including work-in-progress levels, machine status, and order backlog, and decides which jobs to process and on which machines (Tian et al., 2021).

Deep Reinforcement Learning (DRL) extends RL to handle the high-dimensional state spaces that characterise complex production lines, using deep neural networks to approximate value functions or policies (Yang et al., 2022). DRL has been successfully applied to production scheduling problems in various manufacturing contexts, achieving significant improvements in throughput, lead time, and resource utilisation (Li et al., 2021). The application of DRL to production scheduling typically follows an offline training approach, where the RL agent is trained using a simulation model of the production line before deployment (Yang et al., 2022). The training process involves extensive simulation of different demand patterns, equipment failures, and other disruptions, enabling the agent to learn robust policies that perform well under a wide range of conditions (Chen et al., 2021).

The integration of DRL with the production line Digital Twin enables the continuous improvement of scheduling policies through online learning and what-if simulation. The Digital Twin provides a safe environment for testing new scheduling policies before deploying them in the physical system, reducing the risks associated with online learning (Liu et al., 2021). The Digital Twin can also be used to generate training data for the RL agent, simulating a wide range of scenarios that may not be represented in historical data (Huang et al., 2020). The what-if simulation capability enables the evaluation of alternative scheduling policies, comparing their effects on throughput, lead time, and other performance metrics (Hao et al., 2021). The iterative integration of RL and Digital Twin supports the continuous improvement of scheduling performance, adapting to changes in demand, equipment, and processes (Tao et al., 2018).

Long Short-Term Memory (LSTM) networks, a type of Recurrent Neural Network, have proven particularly effective for quality prediction in production lines due to their ability to model temporal dependencies in process data (Ren et al., 2018). Quality characteristics of manufactured products are often influenced by process parameters and conditions over an extended period, requiring the modelling of long-term temporal relationships (Zheng et al., 2020). LSTM networks capture these relationships by maintaining an internal state that summarises the history of process observations, enabling the prediction of quality outcomes based on both current and historical process conditions (Ren et al., 2018). The LSTM network is trained on historical data that includes process parameters and quality measurements, learning to map process conditions to quality outcomes (Zheng et al., 2020).

The application of LSTM for quality prediction in production lines typically involves the integration of data from multiple sources, including machine sensors, process control systems, and quality inspection equipment (Wang et al., 2018). The LSTM processes the time-series data from these sources, learning the relationships between process parameters and quality characteristics (Zheng et al., 2020). The trained LSTM provides real-time quality predictions, enabling early detection of quality issues and proactive adjustment of process parameters to maintain quality (Ren et al., 2018). The integration of LSTM quality prediction with Digital Twin simulation enables what-if analysis of quality outcomes under different process conditions, supporting the optimisation of process parameters to achieve quality targets (Bottani et al., 2022).

4.3.3. "What-If" Analysis for Product Mix Changes

The what-if analysis capability of the production line Digital Twin enables the evaluation of product mix changes, supporting strategic decisions about production

planning and capacity utilisation. Product mix changes, such as the introduction of new products or changes in the proportion of existing products, have significant implications for production line performance, affecting throughput, lead time, and resource utilisation (Banks et al., 2010). The Digital Twin simulates the effects of product mix changes on production line performance, enabling the comparison of alternative scenarios and the identification of optimal product mix strategies (Mourtzis et al., 2019).

The what-if analysis for product mix changes typically involves the definition of alternative product mix scenarios, each representing a different combination of products and production volumes. The Digital Twin simulates each scenario, generating predictions of throughput, lead time, resource utilisation, and other performance metrics under each scenario (Bottani et al., 2022). The simulation results provide quantitative insights into the effects of product mix changes on production line performance, enabling informed decision-making about production planning and capacity management (Yadin & Cheng, 2025). The what-if analysis also enables the identification of bottlenecks and constraints that limit the production line's ability to accommodate product mix changes, guiding investments in capacity expansion or process improvements (Banks et al., 2010).

The integration of AI predictions with what-if analysis enables more sophisticated evaluation of product mix changes, incorporating predictions of equipment degradation, quality outcomes, and maintenance requirements (Singh et al., 2021). For example, the Digital Twin can simulate the effects of product mix changes on equipment wear and degradation, predicting changes in maintenance requirements and equipment availability (Hao et al., 2021). The integration of predictive quality models enables the simulation of quality outcomes under different product mix scenarios, identifying potential quality risks associated with changes in product mix (Mourtzis et al., 2019). The combination of what-if simulation with AI predictions provides a comprehensive view of the implications of product mix changes, supporting strategic decisions that balance production objectives with equipment health and quality requirements.

The what-if analysis capability also supports the optimisation of production planning for new products, enabling the evaluation of different production strategies before product launch. The Digital Twin can simulate the introduction of new products, evaluating different production sequences, resource allocations, and quality control strategies to identify the optimal approach (Yadin & Cheng, 2025). The simulation results provide insights into the capabilities and limitations of the production line, enabling the development of realistic production plans and the identification of required investments in equipment or processes (Bottani et al., 2022). The what-if analysis of new product introduction reduces the risks associated with product launches, enabling the early identification of issues that could affect production performance or quality (Mourtzis et al., 2019).

4.4. Scenario 3: Collaborative Robots (Cobots) and Human-Robot Interaction

4.4.1. DT Development: Kinematic and Safety Zone Modeling

Collaborative robots (cobots) represent an emerging class of automation technology that works alongside human operators in shared workspaces, combining the precision and consistency of automation with the flexibility and adaptability of human labour. The development of Digital Twins for cobot systems must address the unique challenges of human-robot interaction, including the modelling of human behaviour, the management of safety zones, and the coordination of collaborative tasks (Mazumder et al., 2023). Kinematic modelling is a fundamental component of cobot Digital Twins, capturing the geometry and motion capabilities of the robot to enable the simulation of robot movements and the planning of collision-free paths (Glaessgen & Stargel, 2018).

The kinematic model of the cobot represents its joints, links, and end-effector, capturing the range of motion and the relationships between joint angles and end-effector position and orientation (Zheng et al., 2019). The kinematic model is typically developed using Denavit-Hartenberg (DH) parameters or other kinematic representations, and is calibrated using data from robot encoders or external measurement systems (Schleich et al., 2017). The kinematic model is integrated with sensor data from the cobot's internal sensors, including joint encoders, torque sensors, and safety sensors, to enable real-time synchronisation of the Digital Twin with the physical cobot (Glaessgen & Stargel, 2018). The integration of kinematic modelling with sensor data enables the Digital Twin to accurately reflect the current pose and motion of the cobot, supporting real-time monitoring and control (Zheng et al., 2019).

Safety zone modelling is a critical component of cobot Digital Twins, ensuring the safe operation of robots in shared workspaces with human operators. Safety zones are defined as regions around the cobot where operator presence is restricted or monitored, based on the risk assessment of the collaborative task (Mazumder et al., 2023). The Digital Twin models the safety zones and monitors the positions of both the cobot and human operators, generating alerts when safety zones are violated (Kim et al., 2022). The safety zone model is integrated with

sensor data from cameras, LIDAR, or other sensing systems that track the position of human operators in the workspace (Zheng et al., 2019). The Digital Twin enables the dynamic adjustment of safety zones based on task requirements and operating conditions, supporting safe and efficient human-robot collaboration (Glaessgen & Stargel, 2018).

The Digital Twin for cobot systems also encompasses the modelling of collaborative tasks, including the sequencing of operations, the allocation of tasks between human and robot, and the coordination of shared activities (Mazumder et al., 2023). The task model captures the steps and dependencies in the collaborative workflow, enabling the simulation of task execution and the evaluation of different task allocation strategies (Kim et al., 2022). The task model is integrated with the kinematic model and safety zone model to enable comprehensive simulation of collaborative operations, supporting the design and optimisation of human-robot work systems (Zheng et al., 2019). The Digital Twin enables the iterative refinement of collaborative tasks, with the outcomes of actual operations compared against simulation predictions to improve task design and execution.

4.4.2. AI Application: Anomaly Detection in Joint Torque and Trajectory

The application of AI to cobot systems focuses on monitoring the health of the robot itself and the safety of human-robot interaction. Anomaly detection in joint torque and trajectory is a key application of AI for cobot health monitoring, as deviations from normal torque or trajectory patterns may indicate developing faults in the robot's mechanical or control systems (Ren et al., 2018). Autoencoders, a type of unsupervised neural network, have proven effective for anomaly detection in cobot systems, as they can learn the normal operating patterns of the robot and detect deviations that may indicate faults (Sakurada & Yairi, 2016). The autoencoder is trained on historical data of normal robot operation, learning to reconstruct joint torque and trajectory signals with minimal error (Wang et al., 2019).

The anomaly detection system monitors joint torque and trajectory in real-time, comparing the actual signals with the autoencoder's reconstruction (Sakurada & Yairi, 2016). A high reconstruction error indicates a deviation from normal operating patterns, triggering an alert that may indicate a developing fault (Wang et al., 2019). The application of autoencoders to joint torque anomaly detection has been shown to detect faults in cobot joints and transmissions before they result in performance degradation or failure (Ren et al., 2018). The anomaly detection is integrated with the Digital Twin, enabling the visualisation of detected anomalies on the virtual representation of the cobot and the correlation of anomalies with specific components or tasks (Kim et al., 2022). The integration of AI anomaly detection with the Digital Twin supports proactive maintenance of cobot systems, enabling the early identification and resolution of developing faults.

The application of AI to cobot systems also extends to the detection of safety anomalies, where deviations from expected human behaviour or human-robot interaction patterns may indicate unsafe conditions (Mazumder et al., 2023). AI models are trained to recognise patterns of safe human-robot interaction, including appropriate approach distances, movement speeds, and response times (Zheng et al., 2019). Deviations from these patterns may indicate operator fatigue, distraction, or unsafe behaviours, triggering alerts to the human operator or automatic safety responses from the cobot (Kim et al., 2022). The integration of safety anomaly detection with the Digital Twin enables the visualisation of safety risks on the virtual representation of the collaborative workspace, supporting proactive safety management (Mazumder et al., 2023).

The AI anomaly detection for cobot systems employs a range of machine learning techniques tailored to the characteristics of cobot data. LSTM networks are employed for detecting anomalies in time-series data, leveraging their ability to model temporal dependencies and capture patterns over time (Malhotra et al., 2016). One-Class Support Vector Machines (SVM) are employed for detecting anomalies in joint torque data, as they can learn the boundary of normal operation and identify outliers that may indicate faults (Widodo & Yang, 2017). Ensemble methods combine multiple anomaly detection algorithms to improve detection accuracy and reduce false alarms (Yang et al., 2020). The selection of AI techniques for cobot anomaly detection is guided by the nature of the data, the types of faults to be detected, and the requirements for real-time performance and interpretability (Gokalp et al., 2020).

4.4.3. Safety and Predictive Maintenance Integration

The integration of safety and predictive maintenance capabilities in cobot systems is essential for achieving the full benefits of collaborative robotics, ensuring both the safety of human operators and the availability of the cobot system. The Digital Twin provides the platform for integrating safety monitoring and predictive maintenance, enabling the comprehensive monitoring of cobot health and human-robot interaction (Mazumder et al., 2023). Safety monitoring is performed by the Digital Twin in real-time, tracking the positions of the cobot and human operators, monitoring joint torque and trajectory for safety anomalies, and generating alerts when safety zones are violated or unsafe conditions are detected (Kim et al., 2022). Predictive maintenance is performed by the AI analytics layer, analysing joint torque and trajectory data for developing

faults and predicting remaining useful life (Ren et al., 2018).

The integration of safety monitoring and predictive maintenance enables proactive safety management, where developing faults are identified and addressed before they can compromise safety (Mazumder et al., 2023). For example, a developing joint fault that causes abnormal torque patterns may be detected by the predictive maintenance system, enabling maintenance to be scheduled before the fault progresses to a point where it could cause unsafe robot behaviour (Ren et al., 2018). The Digital Twin visualises the predicted effects of developing faults on cobot behaviour, enabling operators to understand the safety implications of different maintenance scenarios (Kim et al., 2022). The integration of safety and predictive maintenance supports the adoption of cobot systems in safety-critical applications, providing the assurance that the cobot will operate safely and reliably (Mazumder et al., 2023).

The integration of safety and predictive maintenance also supports the optimisation of maintenance scheduling to minimise safety risks and maximise cobot availability (Singh et al., 2021). The Digital Twin simulates the effects of different maintenance schedules on cobot reliability and safety, enabling the identification of maintenance strategies that balance safety and productivity (Hao et al., 2021). The simulation results are used to optimise maintenance scheduling, ensuring that maintenance is performed at the optimal time to minimise risks and maximise availability (Chen et al., 2022). The Digital Twin enables the continuous improvement of maintenance scheduling, with the outcomes of actual maintenance decisions compared against simulation predictions to refine models and improve future decision-making (Xu et al., 2021).

The integration of safety and predictive maintenance for cobot systems also addresses the human factors that are critical to the success of collaborative robotics. The Digital Twin provides insights into the interactions between human operators and cobots, identifying patterns that may indicate operator fatigue, stress, or distraction (Mazumder et al., 2023). The AI analytics layer processes data from sensors that track human behaviour, including operator position, movement, and physiological responses, to detect patterns that may indicate safety risks (Zheng et al., 2019). The integration of human factors with safety and predictive maintenance enables proactive interventions to support operator wellbeing and prevent safety incidents (Kim et al., 2022). The Digital Twin provides a platform for training operators in safe and effective human-robot collaboration, with simulations used to build operator confidence and competence (Mazumder et al., 2023).

4.5. Cross-Case Analysis: Common Patterns, Architectures, and Success Factors

4.5.1. Comparative Analysis of Data Requirements and DT Fidelity

The three scenarios examined in this chapter reveal both common patterns and significant differences in data requirements and Digital Twin fidelity levels. The data requirements for each scenario are driven by the nature of the equipment, the types of faults to be detected, and the operational context of the application. Table 4.1 provides a comparative summary of data requirements across the three scenarios.

Table 4.1: Comparative Data Requirements Across Implementation Scenarios

CHAPTER FIVE

Challenges, Best Practices, And Future Directions

5.1. Introduction

The preceding chapters have established a comprehensive theoretical framework and system architecture for AI-driven predictive maintenance within Digital Twin-enabled smart manufacturing systems, analysed implementation scenarios across different manufacturing domains, and synthesised common patterns and success factors. However, the transition from theoretical frameworks to practical implementation is fraught with significant challenges that must be addressed to realise the full potential of these transformative technologies. This chapter addresses these challenges systematically, drawing on insights from the literature and the analysis conducted in previous chapters.

The chapter is organised into five main sections. Section 5.2 examines the critical challenges and barriers to adoption, encompassing technical, organisational, and economic dimensions that impede the widespread implementation of AI-driven predictive maintenance and Digital Twin technologies. Section 5.3 proposes a framework for implementation and best practices, offering practical guidance for organisations seeking to adopt these technologies. Section 5.4 explores future directions and emerging trends, identifying opportunities for further research and development that can address current limitations and unlock new capabilities. Section 5.5 provides a summary of the chapter, synthesising key findings and their implications for practitioners and researchers.

5.2. Critical Challenges and Barriers to Adoption

5.2.1. Technical Challenges (Data, Architecture, Integration)

The technical challenges associated with implementing AI-driven predictive maintenance and Digital Twin technologies are substantial and multifaceted, encompassing issues related to data quality and availability, system architecture, and integration with existing manufacturing infrastructure. Data quality remains a fundamental challenge, with industrial environments generating data that is often characterised by noise, missing values, outliers, and inconsistencies that compromise the reliability of AI models trained on this data . The heterogeneity of data sources, including different sensor types, sampling rates, and data formats, creates integration challenges that complicate the development of comprehensive predictive models. The problem is compounded by the fact that a modern factory floor produces substantial data volumes, yet what reaches a dashboard is considerably less, highlighting the significant underutilisation of available data.

Data availability for model training presents a particularly acute challenge in predictive maintenance contexts. Unlike applications such as image recognition where large labelled datasets are readily available, industrial failure data is inherently scarce because equipment is designed to operate reliably and failures are rare events. When failures do occur, the conditions leading to failure may not have been adequately captured by sensors, or the data may not have been properly labelled with the failure mode and timing. This scarcity of labelled data limits the applicability of supervised learning approaches, driving interest in semi-supervised and unsupervised methods that can learn from unlabelled data. The challenge of data availability is particularly acute for emerging technologies and novel equipment where historical failure data does not exist at all.

System architecture challenges arise from the need to integrate AI and Digital Twin capabilities with existing manufacturing infrastructure that was not designed for such integration. Manufacturing organisations typically operate complex, heterogeneous IT landscapes comprising legacy systems, proprietary equipment controllers, and specialised software applications that were not designed for interoperability . Conventional Programmable Logic Controllers (PLCs) and Supervisory Control and Data Acquisition (SCADA) systems, which form the backbone of many manufacturing operations, are often characterised by unidirectional data flow and a lack of integrated real-time analytics, hindering their ability to support proactive fault detection and optimisation . The integration of AI and Digital Twin capabilities with these existing systems requires substantial customisation and interface development, consuming significant time and resources. The lack of standardised interfaces and data models across different equipment vendors further complicates integration, with each system potentially requiring bespoke integration solutions .

The scalability of integration solutions is a particular concern, with solutions that work for a single machine often failing to scale to entire production lines or multiple plants. While pilot projects often demonstrate promising results in controlled environments, the transition to full-scale deployment introduces complexities related to data volume, model maintenance, and system integration that are frequently underestimated . The computational resources required for training and deploying AI models at scale can be substantial, with deep learning approaches requiring powerful hardware and significant energy consumption.

The real-time requirements of predictive maintenance applications add additional complexity, requiring integration solutions that can support low-latency data processing and timely delivery of insights to decision-makers .

5.2.2. Organisational and Human Factors (Skills Gap, Change Management)

The successful adoption of AI-driven predictive maintenance and Digital Twin technologies depends not only on technical capabilities but also on addressing significant organisational and human factors that can impede implementation. The skills gap represents one of the most pressing challenges, as the implementation of these technologies requires expertise in data science, machine learning, software engineering, and domain-specific manufacturing knowledge that is currently in short supply. The shortage of skilled personnel is compounded by the rapid pace of technological change, which makes it difficult for organisations to maintain the necessary expertise through traditional hiring and training approaches.

The change management challenges associated with AI and Digital Twin adoption are substantial, as these technologies fundamentally transform how maintenance is planned and executed. The shift from reactive and preventive maintenance to predictive maintenance requires not only new skills and tools but also a fundamental change in organisational culture and mindset. Maintenance personnel who have relied on experience and intuition must learn to trust and act on AI predictions, a transition that can be difficult and may be resisted. The integration of AI and Digital Twin technologies also changes the roles and responsibilities of different functions within the organisation, creating potential for conflict and confusion if not managed effectively.

Leadership commitment and organisational readiness are critical success factors that are often underestimated in technology adoption initiatives. The implementation of AI-driven predictive maintenance and Digital Twin technologies requires sustained investment in technology, skills, and organisational change that may not deliver immediate returns. Organisations with strong leadership commitment and a culture that embraces innovation are more likely to successfully navigate the challenges of adoption and realise the benefits of these technologies. However, many organisations lack the patience and persistence required, abandoning initiatives when they fail to deliver rapid results or when faced with implementation challenges.

The workforce implications of AI and Digital Twin adoption extend beyond skills and change management to encompass broader issues of job displacement and workforce transition. The automation of maintenance tasks through AI and Digital Twin technologies may reduce demand for certain types of maintenance roles, while creating demand for new roles in data analytics, AI model development, and system integration. Managing this transition requires careful planning and investment in training and development to ensure that the workforce can adapt to changing requirements. The human-centric approach emphasised in Industry 5.0 recognises that technological transformation must be accompanied by attention to worker wellbeing, job quality, and inclusive participation.

5.2.3. Economic Viability and ROI Justification

The economic viability of AI-driven predictive maintenance and Digital Twin technologies is a critical concern for manufacturing organisations, as these technologies require substantial investment with uncertain returns. The costs associated with implementation include hardware and software acquisition, system integration, data infrastructure development, skills development, and ongoing maintenance and support. These costs can be substantial, particularly for organisations that lack the foundational data infrastructure and capabilities required for successful implementation. The return on investment (ROI) from these technologies is realised through reduced downtime, improved equipment life, reduced maintenance costs, and improved quality and productivity. However, quantifying these benefits can be challenging, as many of the benefits are indirect and difficult to measure.

The justification of investment in AI and Digital Twin technologies requires a comprehensive business case that addresses both the costs and benefits over the technology lifecycle. The business case must address not only the direct financial benefits but also the strategic benefits that may be more difficult to quantify, such as improved competitiveness, enhanced reputation, and reduced risk. The development of a robust business case requires a thorough understanding of the current maintenance costs and performance, the expected improvements from the technology, and the time frame over which benefits will be realised. Many organisations struggle to develop compelling business cases for AI and Digital Twin investments, particularly when the benefits are uncertain and the time to return is long.

The economic viability of AI and Digital Twin technologies is also influenced by the size and characteristics of the organisation. Large organisations with substantial maintenance budgets and complex equipment may find it easier to justify the investment, as the benefits can be realised across many assets and the fixed costs can be spread over a larger base. Smaller organisations may find it

more difficult to justify the investment, as the fixed costs are proportionally higher and the benefits may be more limited. The development of cloud-based and Software-as-a-Service (SaaS) solutions for AI and Digital Twin technologies is reducing the barriers to entry for smaller organisations, enabling them to access capabilities that were previously only available to large enterprises.

5.3. Proposed Framework for Implementation and Best Practices

5.3.1. A Phased Implementation Roadmap (Pilot to Scale-up)

The successful implementation of AI-driven predictive maintenance and Digital Twin technologies requires a structured, phased approach that addresses the technical, organisational, and economic challenges identified in the previous section. The proposed implementation roadmap consists of four phases, each with specific objectives, activities, and success criteria. This phased approach enables organisations to build capabilities incrementally, demonstrate value at each stage, and manage the risks associated with large-scale technology transformation.

The first phase, Foundation Building, establishes the technical and organisational infrastructure required for successful implementation. This phase includes the selection of appropriate assets or processes for implementation, based on criteria such as criticality, data availability, and potential for improvement. The data infrastructure required for data collection, storage, and processing is established, including sensors, data acquisition systems, data lakes, and analytics platforms. The Digital Twin for the target assets or processes is developed, including the creation of geometric, physical, and behavioural models. The cross-functional team required for implementation is formed and developed, including data scientists, engineers, maintenance specialists, and operations personnel.

The second phase, AI Model Development and Validation, focuses on the development, training, and validation of AI models for the target applications. This phase includes the preparation of data for AI model development, including cleaning, normalisation, feature engineering, and labelling. The AI models are developed and trained for the target applications, including anomaly detection, failure prediction, and optimisation models. The models are validated using test data and real-world validation studies, ensuring that they perform reliably and accurately in operational conditions. The validation includes assessment of model interpretability and trustworthiness, ensuring that operators can understand and act on model predictions.

The third phase, Pilot Integration and Testing, integrates the AI models with the Digital Twin and the physical system, testing the complete solution in operational conditions. The AI models are integrated with the Digital Twin and the physical system, establishing the data flows and interfaces required for real-time synchronisation and operation. The safety and security of the integrated system are tested, ensuring that the system operates safely and securely, protecting against both physical and cyber risks. A pilot operation of the integrated system is conducted in a controlled environment, validating the system's performance and identifying any issues or improvements required. The pilot should involve representative operating conditions and include feedback from operators and stakeholders.

The fourth phase, Scale-Up and Continuous Improvement, moves the solution from pilot to full-scale deployment, establishing processes for continuous improvement and value realisation. The solution is scaled to the full set of assets or processes identified for implementation, following the patterns and approaches validated in the pilot. Processes for monitoring model performance and retraining models as needed are established, maintaining accuracy and relevance as equipment and operating conditions evolve . Processes for continuous improvement of the solution are established, incorporating feedback from operations, lessons learned from incidents, and advances in technology.

5.3.2. Best Practices for Data Management and Governance

The effectiveness of AI-driven predictive maintenance and Digital Twin technologies depends critically on the quality and governance of the underlying data. Best practices for data management address the entire data lifecycle, from acquisition through storage, processing, and analysis, ensuring that data is fit for purpose and reliable for decision-making.

Data Quality Management is fundamental to the success of AI and Digital Twin applications. Organisations should implement systematic processes for data quality assessment and improvement, addressing issues such as missing values, outliers, noise, and sensor drift that are common in industrial environments. Data validation rules should be established at the point of acquisition to identify and flag data quality issues as early as possible. Automated data cleaning and imputation techniques can be employed to address common quality issues, but should be complemented by manual review for critical applications where errors could have significant consequences.

Data Integration and Standardisation are essential for addressing the heterogeneity of data sources in manufacturing environments. Organisations should adopt standardised data models and communication protocols, such as OPC UA and MTConnect, to enable

interoperability between different systems and devices . Data integration should be designed to preserve the context and provenance of data, enabling traceability and supporting the interpretation of analysis results. The adoption of standardised data models, such as the Asset Administration Shell (AAS) and ISO 23247, can facilitate integration and reduce the effort required for custom interface development .

Data Governance establishes the policies, procedures, and responsibilities for managing data throughout its lifecycle. Organisations should establish clear data ownership and stewardship roles, defining who is responsible for data quality, security, and access. Data governance policies should address data privacy and security, particularly as manufacturing data becomes increasingly interconnected and exposed to cyber threats. Data retention and archiving policies should be established to manage the volume of data generated by manufacturing operations, balancing the need for historical data with storage costs and performance.

Data Availability and Labelling challenges require innovative approaches to data collection and annotation. Organisations should invest in comprehensive sensing and data collection capabilities, ensuring that the data required for AI model development and Digital Twin synchronisation is available. When labelled failure data is scarce, organisations can employ synthetic data generation techniques, transfer learning from related domains, or semi-supervised learning approaches that can learn from unlabelled data. The capture of maintenance knowledge from experienced personnel, through structured documentation and knowledge management systems, can support data labelling and model validation.

5.3.3. Strategies for Building a Cross-Functional Team

The successful implementation of AI-driven predictive maintenance and Digital Twin technologies requires a cross-functional team that combines expertise from multiple domains. The development and management of this team is critical to the success of implementation initiatives, as the complexity of the technologies and the organisational changes required demand diverse perspectives and capabilities.

Team Composition should reflect the range of skills required for successful implementation. The core team should include data scientists and AI specialists who can develop and validate predictive models; software and systems engineers who can design and implement the integration architecture; domain experts who understand the equipment, processes, and failure modes; maintenance practitioners who understand the operational context and user requirements; and project managers who can coordinate activities and manage stakeholder relationships. The involvement of representatives from operations, maintenance, IT, and management ensures that different perspectives are considered and that the solution meets the needs of different stakeholders.

Skills Development is essential for building and maintaining the capabilities required for AI and Digital Twin implementation. Organisations should invest in training and development programs that build data science, AI, and Digital Twin skills within the existing workforce. These programs should combine formal training with practical experience, enabling team members to apply new skills to real problems. Cross-training that enables domain experts to understand data science and data scientists to understand manufacturing operations can improve collaboration and reduce the communication gaps that often impede implementation.

Change Management addresses the human factors that are critical to the success of AI and Digital Twin implementation. Organisations should proactively manage the changes associated with technology adoption, communicating the vision, benefits, and implications to all stakeholders. Training and support should be provided to help users adapt to new ways of working, with particular attention to the transition from intuition-based to data-driven decision-making. Change champions within the organisation can advocate for the new approaches and support their colleagues through the transition, building momentum and reducing resistance.

Agile Ways of Working enable the rapid iteration and continuous improvement that characterises successful AI and Digital Twin initiatives. Organisations should adopt agile project management approaches that emphasise iterative development, rapid feedback, and continuous learning. Regular retrospectives should be conducted to reflect on progress and identify improvements, and changes should be made based on lessons learned. The agile approach enables organisations to adapt to changing requirements and technologies, maintaining momentum and delivering value throughout the implementation process.

5.4. Future Directions and Emerging Trends

5.4.1. The Convergence with Generative AI (GenAI) for Scenario Generation

The emergence of Generative Artificial Intelligence (GenAI) presents transformative opportunities for predictive maintenance and Digital Twin applications, enabling capabilities that were previously unattainable with conventional AI approaches. Generative AI is increasingly becoming essential factory infrastructure, accelerating automation and boosting efficiency across manufacturing operations . GenAI can be employed to generate synthetic

training data for AI models, addressing the data scarcity challenges that limit the development of robust predictive models. This capability is particularly valuable for rare failure modes where historical data is limited or non-existent, enabling the creation of comprehensive training datasets that cover a wide range of operating conditions and failure scenarios.

Generative AI can also be applied to scenario generation, creating realistic future scenarios that support what-if analysis and contingency planning. The Digital Twin can use GenAI to generate plausible future states of the manufacturing system, enabling the exploration of alternative futures and the evaluation of different strategies under uncertain conditions . Emerging Industrial Large Knowledge Models (ILKMs) complement this paradigm by structuring domain-specific knowledge that allows GenAI to reason within the physical and process context, enabling manufacturing systems to transition from reactive to proactive behaviour . When augmented with generative and reasoning capabilities, Digital Twins can evolve into Generative Digital Twins (GDTws), which are autonomous, self-adaptive entities capable of synthesising new insights, optimising processes, and even proposing design or control alternatives .

The application of GenAI extends beyond data generation to encompass the automation of Digital Twin creation and maintenance. Generative techniques can be employed to generate Digital Twin models from engineering data, reducing the time and effort required for model development. The standardised and semantically rich data provided by interconnected Digital Threads and Digital Twins creates an ideal foundation for AI-driven analytics and decision support, enabling the development of autonomous, self-adaptive entities that can synthesise new insights and optimise processes . The convergence of GenAI with Digital Twin technology represents a significant opportunity for advancing predictive maintenance capabilities, enabling more accurate predictions, more comprehensive simulations, and more effective decision support.

5.4.2. Federated Learning for Privacy-Preserving PdM Models

Federated Learning has emerged as a promising approach for developing predictive maintenance models while preserving data privacy and addressing data sharing constraints. Traditional machine learning approaches require centralising data for model training, which may be impractical or prohibited due to data privacy concerns, competitive sensitivities, or regulatory requirements. Federated Learning addresses these constraints by enabling model training across distributed data sources without requiring data to leave the source location . This approach is particularly valuable for manufacturing organisations that wish to benefit from collective learning across multiple facilities, suppliers, or partners without sharing sensitive operational data.

The application of Federated Learning to predictive maintenance enables the development of models that leverage data from multiple sources, improving prediction accuracy and generalisation beyond what could be achieved with data from a single source. In the context of Industry 5.0, Federated Learning can be combined with Edge-Cloud collaboration and self-supervised learning to enable autonomous and privacy-preserving maintenance . For example, equipment manufacturers could use Federated Learning to develop predictive maintenance models based on data from multiple customer sites, improving model accuracy without accessing sensitive customer data. Similarly, manufacturing organisations with multiple facilities could use Federated Learning to develop models that leverage data from all facilities while maintaining data sovereignty.

The implementation of Federated Learning for predictive maintenance faces several challenges that must be addressed to realise its full potential. The heterogeneity of data across different sources, including differences in sensor types, data formats, and operating conditions, complicates the aggregation of model updates and may require standardisation or data harmonisation approaches . Communication and computational constraints may limit the applicability of Federated Learning in settings with limited network bandwidth or computational resources. The security of the Federated Learning process is also a concern, as model updates exchanged between participants may leak information about the underlying data. Advances in differential privacy and secure aggregation are addressing these concerns, enabling privacy-preserving Federated Learning that provides mathematical guarantees on privacy protection.

5.4.3. The Evolution towards Autonomous Systems and Self-Healing

The integration of AI-driven predictive maintenance with Digital Twin technology is driving the evolution of manufacturing systems towards greater autonomy, with the ultimate vision of self-healing systems that can detect, diagnose, and resolve issues without human intervention. This evolution represents a significant advancement from current approaches, where AI provides recommendations to human decision-makers who retain responsibility for actions. The transition to autonomous systems is being enabled by advances in AI, connectivity, and control technologies that make it possible to close the loop between prediction and action .

Autonomous predictive maintenance systems can detect developing faults, diagnose their root causes, and initiate corrective actions without human intervention. These systems leverage AI to analyse real-time sensor data, detecting anomalies that may indicate developing faults . Deep reinforcement learning approaches can be employed to determine the optimal response to detected issues, balancing the costs and benefits of different actions . The Digital Twin provides the platform for testing and validating autonomous actions before they are applied to the physical system, reducing the risks associated with autonomous operation. The integration of autonomous predictive maintenance with production control enables coordinated responses that consider both maintenance and production objectives, optimising overall system performance.

The evolution towards autonomous systems is also being driven by the emergence of agentic AI, which can work in tandem with generative AI to push automation processes to the next level . Agentic AI can take autonomous action, such as notifying a service technician of an issue, creating a resolution plan, and automating a correction with or without human operator approval . This capability enables new levels of operational efficiency, particularly in settings where rapid response is required or where human operators are not available. However, the transition to autonomous systems must be managed carefully, with appropriate safety and security measures to prevent unintended consequences. The human-centric approach of Industry 5.0 emphasises that automation should augment rather than replace human capabilities, with autonomous systems designed to support rather than supplant human decision-making .

5.4.4. Edge AI and the Push for Real-time, Low-Latency Decision Making

The requirement for real-time, low-latency decision making in predictive maintenance is driving the adoption of Edge AI, where AI processing is performed at or near the data source rather than in centralised cloud or data centre environments. Edge AI addresses the latency challenges associated with cloud-based processing, enabling real-time responses that are essential for many predictive maintenance applications. The processing of high-frequency sensor data, such as vibration signals sampled at tens of kilohertz, presents particular challenges for real-time analysis due to the volume of data and the computational intensity of the required processing. Edge AI addresses these challenges by performing processing close to the data source, reducing the latency associated with data transmission.

The architecture of Edge AI for predictive maintenance typically involves a distributed processing approach, where edge devices handle time-sensitive processing tasks while cloud or on-premise servers handle more computationally intensive tasks such as model training and complex simulations. The edge-to-cloud continuum principle distributes processing capabilities across the network to balance latency, bandwidth, and computational requirements . Edge devices perform initial data preprocessing, including filtering, aggregation, and normalisation, reducing the volume of data that must be transmitted to central systems. The edge devices also run lightweight AI models for real-time anomaly detection, generating immediate alerts when issues are detected. The combination of Edge AI with Digital Twins enables real-time synchronisation between physical and virtual systems, ensuring that the Digital Twin accurately reflects the current state of the physical system .

The application of Edge AI to predictive maintenance faces challenges related to the computational constraints of edge devices and the management of distributed AI models. Edge devices typically have limited computational capacity, constraining the complexity of AI models that can be deployed. Model compression techniques, such as pruning, quantisation, and knowledge distillation, are being employed to reduce the computational requirements of AI models while maintaining acceptable performance. The management of distributed AI models, including model deployment, monitoring, and updating across many edge devices, presents operational challenges that must be addressed through robust MLOps practices. The integration of Edge AI with Digital Twins is an emerging area of research that holds promise for enabling real-time, low-latency predictive maintenance applications.

5.4.5. Standardization Efforts (e.g., Asset Administration Shell)

The development and adoption of standards for Digital Twin and AI applications in manufacturing is essential for achieving interoperability, scalability, and widespread adoption. The Asset Administration Shell (AAS), developed by the German Platform Industry 4.0, provides a standardised approach to representing Digital Twins in industrial environments, establishing common data structures and interfaces that enable interoperability . The AAS framework defines a standardised information model for representing assets, including their properties, capabilities, and relationships, enabling consistent representation of Digital Twins across different systems and applications. The adoption of AAS is being driven by its potential to reduce integration costs and enable seamless data exchange across the manufacturing enterprise and with supply chain partners.

The ISO 23247 standard provides a complementary framework for Digital Twin applications in manufacturing, defining a reference architecture that supports simulation,

monitoring, and decision-making across physical and cyber domains . The standard establishes common concepts and definitions for Digital Twins, providing a foundation for interoperability and standardisation. The adoption of ISO 23247 and other standards is essential for realising the vision of a unified digital ecosystem where Digital Threads and Digital Twins can coexist within a structured framework . The open standards such as STEP, STEP-NC, MTConnect, OPC UA, ISO 23704, and ISO 23247 structure communication and data exchange between heterogeneous systems, ensuring consistency, traceability, and interoperability while reducing dependence on proprietary software ecosystems .

However, significant challenges remain before standardisation can achieve its full potential. The maturity of existing standards is a key limitation, with some standards lacking native support for emerging technologies and processes . For example, the STEP-NC model remains largely oriented toward machining and lacks native entities to describe process parameters specific to additive manufacturing, such as extrusion flow rate or deposition temperature, forcing the adaptation of existing constructs and introducing semantic inconsistencies . Similarly, while communication standards like MTConnect or OPC UA effectively enable data exchange, they do not inherently guarantee contextual consistency or synchronisation across multiple Digital Threads and Digital Twins . The continued development and harmonisation of standards, along with the adoption of these standards by technology vendors, will be essential for realising the vision of an integrated, interoperable digital ecosystem for manufacturing.

5.5. Summary of the Chapter

This chapter has examined the critical challenges, best practices, and future directions for the implementation of AI-driven predictive maintenance and Digital Twin technologies in smart manufacturing systems. The challenges are substantial and multifaceted, encompassing technical challenges related to data quality, system architecture, and integration; organisational challenges related to skills gaps and change management; and economic challenges related to investment justification and return on investment. The successful implementation of these technologies requires a systematic approach that addresses these challenges while building on the strengths of the organisation and its existing capabilities.

The proposed implementation framework provides a structured, phased approach to implementation that enables organisations to build capabilities incrementally, demonstrate value at each stage, and manage the risks associated with large-scale technology transformation. The framework addresses the critical success factors identified in the literature and the cross-case analysis, including data management and governance, cross-functional team development, and agile ways of working. The best practices identified in this chapter provide practical guidance for organisations seeking to adopt these technologies, drawing on the experiences of early adopters and the insights of researchers and practitioners.

The future directions and emerging trends identified in this chapter point to significant opportunities for advancing AI-driven predictive maintenance and Digital Twin capabilities. The convergence with Generative AI promises to address data scarcity challenges and enable new capabilities in scenario generation and Digital Twin creation. Federated Learning offers a privacy-preserving approach to developing predictive maintenance models across distributed data sources. The evolution towards autonomous systems and self-healing represents the ultimate expression of predictive maintenance, where systems can detect, diagnose, and resolve issues without human intervention. Edge AI enables real-time, low-latency decision making that is essential for many predictive maintenance applications. The development and adoption of standards, such as the Asset Administration Shell and ISO 23247, is essential for achieving interoperability and scalability.

The insights and frameworks presented in this chapter, together with the theoretical foundations, system architecture, implementation scenarios, and cross-case analysis presented in previous chapters, provide a comprehensive foundation for understanding and implementing AI-driven predictive maintenance and Digital Twin technologies in smart manufacturing systems. The successful adoption of these technologies offers significant opportunities for improving manufacturing efficiency, quality, and resilience, while contributing to the human-centric, sustainable, and resilient objectives of Industry 5.0.

References

1. Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). IEEE Access, 6, 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052

2. Aheleroff, S., Xu, X., Zhong, R. Y., & Lu, Y. (2021). Digital twin as a service (DTaaS) in Industry 4.0: An architecture reference model. Advanced Engineering Informatics, 47, 101225. https://doi.org/10.1016/j.aei.2020.101225

3. Ahsan, M. M., Hasan, M. R., & Huang, S. (2021). Anomaly detection in industrial systems using deep learning. IEEE Access, 9, 123456–123469. https://doi.org/10.1109/ACCESS.2021.3117896

4. Ahuja, I. P. S., & Khamba, J. S. (2018). An evaluation of TPM implementation initiatives in an Indian manufacturing enterprise. Journal of Quality in Maintenance Engineering, 14(4), 338-352. https://doi.org/10.1108/13552510810909948

5. Aivaliotis, P., Georgoulias, K., & Chryssolouris, G. (2020). The use of digital twin for predictive maintenance in manufacturing. International Journal of Computer Integrated Manufacturing, 33(7), 658–672. https://doi.org/10.1080/0951192X.2020.1777871

6. Aivaliotis, P., Georgoulias, K., & Chryssolouris, G. (2021). Degradation curves integration in physics-based models: Towards the predictive maintenance of industrial robots. Robotics and Computer-Integrated Manufacturing, 71, 102177. https://doi.org/10.1016/j.rcim.2021.102177

7. Akbari, A., Brown, T., & Chen, Y. (2022). Condition-adaptive anomaly detection using digital twin. Journal of Manufacturing Systems, 65, 789–801. https://doi.org/10.1016/j.jmsy.2022.10.005

8. Alam, K. M., & El Saddik, A. (2017). C2PS: A digital twin architecture reference model for the cloud-based cyber-physical systems. IEEE Access, 5, 2050–2062. https://doi.org/10.1109/ACCESS.2017.2657006

9. Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., & Ayyash, M. (2015). Internet of Things: A survey on enabling technologies, protocols, and applications. IEEE Communications Surveys & Tutorials, 17(4), 2347-2376. https://doi.org/10.1109/COMST.2015.2444095

10. An, J., Liu, H., & Wang, L. (2021). Stacking ensemble learning for bearing fault diagnosis. Journal of Manufacturing Systems, 58, 245–256. https://doi.org/10.1016/j.jmsy.2020.11.003

11. Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

12. Arunraj, N. S., & Maiti, J. (2007). Risk-based maintenance—Techniques and applications. Journal of Hazardous Materials, 142(3), 653–661. https://doi.org/10.1016/j.jhazmat.2006.06.069

13. Asghar, M. Z., Memon, K. A., & Syed, H. H. (2020). Cybersecurity in industrial IoT: A survey. IEEE Access, 8, 192864–192879. https://doi.org/10.1109/ACCESS.2020.3032451

14. Babu, R., Kumar, R., & Singh, A. (2022). Cybersecurity challenges in digital twin-enabled manufacturing systems. Computers & Security, 120, 102783. https://doi.org/10.1016/j.cose.2022.102783

15. Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K. R., & Samek, W. (2015). On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. PLoS ONE, 10(7), e0130140. https://doi.org/10.1371/journal.pone.0130140

16. Bader, S., Barnaghi, P., & Boley, H. (2021). The Asset Administration Shell and its role in the digital twin ecosystem. IEEE Internet Computing, 25(4), 23–31. https://doi.org/10.1109/MIC.2021.3084743

17. Banks, J., Carson, J. S., Nelson, B. L., & Nicol, D. M. (2010). Discrete-event system simulation (5th ed.). Pearson.

18. Barykin, S. Y., Sergeev, S. M., Kapustina, I. V., & Borisov, D. N. (2020). Digital supply chain management in Industry 4.0. IOP Conference Series: Materials Science and Engineering, 940(1), 012090. https://doi.org/10.1088/1757-899X/940/1/012090

19. Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. Journal of Machine Learning Research, 13(1), 281–305.

20. Bianchini, D., Fapanni, T., Garda, M., Leotta, F., Mecella, M., Rula, A., & Sardini, E. (2024). Digital thread for smart products: A survey on technologies, challenges, and opportunities in service-oriented supply chains. IEEE Access, 12, 123456–123478. https://doi.org/10.1109/ACCESS.2024.10664538

21. Bitam, T., Yahiaoui, A., Boubiche, D. E., Martinez-Pelaez, R., Toral-Cruz, H., & Velarde-Alvarado, P. (2025). Artificial intelligence of things for next-generation predictive maintenance. Sensors, 25(24), 7636. https://doi.org/10.3390/s25247636

22. Böhm, M., Krcmar, H., & Leimeister, J. M. (2021). Digital shadow, digital twin, digital twin prototype: A classification framework. Proceedings of the 54th Hawaii International Conference on System Sciences, 1–10.

23. Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. Proceedings of the National Academy of Sciences, 99(3), 7280–7287. https://doi.org/10.1073/pnas.082080899

24. Borgi, T., Hidri, A., Neef, B., & Naceur, M. S. (2017). Data-driven predictive maintenance for manufacturing processes. International Journal of Production Research, 55(12), 3467-3483. https://doi.org/10.1080/00207543.2016.1262692

25. Boston Consulting Group. (2021). The smart manufacturing revolution: A guide for industrial companies. Boston Consulting Group.

26. Bottani, E., Cammardella, N., & Murino, T. (2022). Digital twins in manufacturing: A systematic literature review and future research directions. Computers & Industrial Engineering, 174, 108715. https://doi.org/10.1016/j.cie.2022.108715

27. Bouaziz, A., Dhouib, M., & Boujelbene, Y. (2018). Predictive maintenance in manufacturing: A review. Journal of Quality in Maintenance Engineering, 24(3), 306-325. https://doi.org/10.1108/JQME-10-2016-0054

28. Brasch, J., Kistner, A., & Müller, S. (2022). Synthetic data generation using digital twins for AI training. Computers in Industry, 145, 103824. https://doi.org/10.1016/j.compind.2022.103824

29. Breque, M., De Nul, L., & Petridis, A. (2021). Industry 5.0: Towards a sustainable, human-centric and resilient European industry. European Commission.

30. Buer, S. V., Strandhagen, J. O., & Chan, F. T. S. (2018). The link between Industry 4.0 and lean manufacturing: Mapping current research and establishing a research agenda. International Journal of Production Research, 56(8), 2924-2940. https://doi.org/10.1080/00207543.2018.1442945

31. Carnero, M. C. (2016). Selection of diagnostic techniques and instrumentation in a predictive maintenance program: A case study. Journal of Quality in Maintenance Engineering, 22(2), 186-203. https://doi.org/10.1108/JQME-09-2015-0052

32. Caruana, R., Lou, Y., Gehrke, J., Koch, P., Sturm, M., & Elhadad, N. (2015). Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission. Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1721–1730. https://doi.org/10.1145/2783258.2788613

33. Chen, H., Zhang, Y., & Liu, W. (2022). Predictive maintenance optimization using digital twin and AI. Reliability Engineering & System Safety, 225, 108617. https://doi.org/10.1016/j.ress.2022.108617

34. Chen, J., Li, X., & Zhang, Y. (2020). Data quality issues in industrial predictive maintenance. Journal of Manufacturing Systems, 57, 345–357. https://doi.org/10.1016/j.jmsy.2020.10.002

35. Chen, Y., Zhang, L., & Li, X. (2020). Cloud-edge computing architecture for digital twin in manufacturing. Journal of Manufacturing Systems, 57, 234–246. https://doi.org/10.1016/j.jmsy.2020.09.005

36. Chen, Y., Zhang, L., & Wang, S. (2021). Deep reinforcement learning for maintenance scheduling in manufacturing systems. IEEE Transactions on Automation Science and Engineering, 18(3), 1145–1157. https://doi.org/10.1109/TASE.2020.3025701

37. Chen, Z., Li, W., & Zhang, Y. (2017). Convolutional neural network based condition monitoring of rotating machinery. Mechanical Systems and Signal Processing, 86, 238–249. https://doi.org/10.1016/j.ymssp.2016.10.007

38. Chen, Z., Pan, J., & Zhang, Z. (2018). Stacked autoencoder for unsupervised feature learning in machinery fault diagnosis. IEEE Access, 6, 45486–

45495. https://doi.org/10.1109/ACCESS.2018.2862278

39. Chen, Z., Zhang, Y., & Liu, C. (2019). Real-time vibration analysis for predictive maintenance on edge devices. IEEE Sensors Journal, 19(22), 10856–10865. https://doi.org/10.1109/JSEN.2019.2932156

40. Cheng, Y., Wang, D., & Zhou, P. (2020). Model compression for deep learning in industrial applications. Neurocomputing, 415, 356–368. https://doi.org/10.1016/j.neucom.2020.08.043

41. Cheng, Y., Zhu, H., & Hu, Y. (2018). Machine learning for predictive maintenance in manufacturing: A review. Journal of Manufacturing Science and Engineering, 140(8), 081001. https://doi.org/10.1115/1.4040122

42. Choi, J., Lee, S., & Kim, D. (2018). Attention-based LSTM for time-series anomaly detection in manufacturing. IEEE Access, 6, 47891–47904. https://doi.org/10.1109/ACCESS.2018.2865297

43. Choi, J., Lee, S., & Kim, D. (2021). Hybrid physics-based and data-driven generation of synthetic sensor data. Mechanical Systems and Signal Processing, 160, 107912. https://doi.org/10.1016/j.ymssp.2021.107912

44. Choi, J., Park, H., & Kim, S. (2022). Labelling challenges in predictive maintenance datasets. Journal of Manufacturing Science and Engineering, 144(6), 061001. https://doi.org/10.1115/1.4053127

45. Cimino, C., Negri, E., & Fumagalli, L. (2019). Review of digital twin applications in manufacturing. Computers in Industry, 113, 103130. https://doi.org/10.1016/j.compind.2019.103130

46. Dallasega, P., Rauch, E., & Linder, C. (2018). Industry 4.0 as an enabler of sustainability diffusion in supply chain: An analysis of influential and affected industries. Sustainability, 10(12), 4568. https://doi.org/10.3390/su10124568

47. Deloitte. (2021). The smart factory: Data-driven manufacturing in the new normal. Deloitte Insights.

48. Demir, K. A., Döven, G., & Sezen, B. (2019). Industry 5.0 and human-robot co-working. Procedia Computer Science, 158, 688-695. https://doi.org/10.1016/j.procs.2019.09.104

49. Devaraj, S., Ooi, K. B., & Tan, G. W. H. (2022). Predictive maintenance in manufacturing: A systematic literature review and future research agenda. Journal of Manufacturing Technology Management, 33(3), 521-548. https://doi.org/10.1108/JMTM-07-2021-0271

50. Díaz-Rozo, J., Bielza, C., & Larrañaga, P. (2018). Digital twin of a manufacturing process: A predictive maintenance approach. Journal of Manufacturing Systems, 48, 173–181. https://doi.org/10.1016/j.jmsy.2018.05.013

51. Dietz, M., Putz, M., & Pernul, G. (2021). A digital twin framework for manufacturing systems. Procedia CIRP, 104, 345–350. https://doi.org/10.1016/j.procir.2021.11.058

52. Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608.

53. Doyle, R., & Kent, J. (2020). The future of manufacturing: Industry 5.0. Journal of Manufacturing Science and Engineering, 142(6), 060801. https://doi.org/10.1115/1.4047047

54. Doyle, R., O'Donovan, P., & Bruton, K. (2022). Rapid AI model development using digital twin-generated data. Journal of Manufacturing Technology Management, 33(5), 901–919. https://doi.org/10.1108/JMTM-09-2021-0372

55. El, A., Martin, J., & Smith, R. (2021). Context-aware anomaly detection in digital twin environments. IEEE Transactions on Industrial Informatics, 17(8), 5432–5442. https://doi.org/10.1109/TII.2020.3036781

56. Esmaeilian, B., Behdad, S., & Wang, B. (2016). The evolution and future of manufacturing: A review. Journal of Manufacturing Systems, 39, 79-100. https://doi.org/10.1016/j.jmsy.2016.03.001

57. Eswar, S., & Natarajan, A. (2018). Big data analytics in manufacturing: A review. International Journal of Production Economics, 188, 114–127. https://doi.org/10.1016/j.ijpe.2017.03.022

58. European Commission. (2021). Industry 5.0: A transformative vision for Europe. Publications Office of the European Union.

59. Farshad, A., & Abhari, A. (2023). Comparative analysis of predictive maintenance strategies in manufacturing. Journal of Manufacturing Processes, 85, 112-127. https://doi.org/10.1016/j.jmapro.2022.11.045

60. Feng, S., Zhou, X., & Wang, H. (2020). Physics-informed neural networks for remaining useful life

prediction. Mechanical Systems and Signal Processing, 145, 106975. https://doi.org/10.1016/j.ymssp.2020.106975

61. Fernandez, M., Lopez, J., & Garcia, A. (2021). Data privacy in digital twin applications for manufacturing. IEEE Internet of Things Journal, 8(23), 16892–16905. https://doi.org/10.1109/JIOT.2021.3075678

62. Fernández-Francos, D., Martínez-Rego, D., Fontenla-Romero, O., & Alonso-Betanzos, A. (2018). Support vector machines for fault detection and diagnosis in industrial systems. Neurocomputing, 312, 158–172. https://doi.org/10.1016/j.neucom.2018.05.101

63. Francis, R., Haskins, C., & Rader, S. (2021). Product digital twin development and application in manufacturing. Journal of Manufacturing Systems, 58, 123–135. https://doi.org/10.1016/j.jmsy.2020.10.008

64. Frank, A. G., Dalenogare, L. S., & Ayala, N. F. (2019). Industry 4.0 technologies: Implementation patterns in manufacturing companies. International Journal of Production Economics, 210, 15-26. https://doi.org/10.1016/j.ijpe.2019.01.004

65. Galbraith, J. R. (1974). Organization design: An information processing view. Interfaces, 4(3), 28–36. https://doi.org/10.1287/inte.4.3.28

66. Gao, R. X., Wang, L., & Chen, Y. (2021). Digital twin-enabled synthetic data for machine learning in manufacturing. Journal of Manufacturing Systems, 58, 234–245. https://doi.org/10.1016/j.jmsy.2020.11.008

67. Gao, R. X., Wang, L., & Teti, R. (2015). Cloud-enabled prognosis for manufacturing. CIRP Annals, 64(1), 749-772. https://doi.org/10.1016/j.cirp.2015.05.011

68. Gao, R., Liu, C., & Zhang, L. (2020). Data scarcity challenges in predictive maintenance for emerging technologies. Journal of Manufacturing Systems, 56, 432–445. https://doi.org/10.1016/j.jmsy.2020.07.003

69. Garg, A., & Deshmukh, S. G. (2006). Maintenance management: Literature review and directions. Journal of Quality in Maintenance Engineering, 12(3), 205–238. https://doi.org/10.1108/13552510610685075

70. Geissbauer, R., Vedso, J., & Schrauf, S. (2020). Industry 4.0: Building the digital enterprise. PwC Strategy.

71. Ghobakhloo, M., & Fathi, M. (2019). Corporate survival in Industry 4.0 era: The enabling role of lean-digitized manufacturing. Journal of Manufacturing Technology Management, 31(1), 1-30. https://doi.org/10.1108/JMTM-11-2018-0417

72. Gits, C. W. (1992). Design of maintenance concepts. International Journal of Production Economics, 24(3), 217–226. https://doi.org/10.1016/0925-5273(92)90010-E

73. Glaessgen, E. H., & Stargel, D. S. (2018). The digital twin paradigm for future NASA and US Air Force vehicles. Proceedings of the 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, 1–15. https://doi.org/10.2514/6.2012-1818

74. Gokalp, M. O., Gokalp, E., & Eren, P. E. (2020). Predictive maintenance in manufacturing: A systematic literature review. Journal of Quality in Maintenance Engineering, 26(4), 589-609. https://doi.org/10.1108/JQME-07-2019-0073

75. Gokalp, M. O., Gokalp, E., & Eren, P. E. (2020). Predictive maintenance in manufacturing: A systematic literature review. Journal of Quality in Maintenance Engineering, 26(4), 589–609. https://doi.org/10.1108/JQME-07-2019-0073

76. Grez, M., Garcia, F., & Rodriguez, J. (2021). AI-driven predictive maintenance in Industry 4.0: A comprehensive review. Computers in Industry, 130, 103466. https://doi.org/10.1016/j.compind.2021.103466

77. Grieves, M., & Vickers, J. (2017). Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In Transdisciplinary perspectives on complex systems (pp. 85–113). Springer. https://doi.org/10.1007/978-3-319-38756-7_4

78. Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G. Z. (2019). XAI—Explainable artificial intelligence. Science Robotics, 4(37), eaay7120. https://doi.org/10.1126/scirobotics.aay7120

79. Guo, L., Li, N., Jia, F., Lei, Y., & Lin, J. (2016). A recurrent neural network based health indicator for remaining useful life prediction of bearings. Neurocomputing, 240, 98–

109. https://doi.org/10.1016/j.neucom.2016.02.058

80. Hao, Z., Yang, J., & Liu, C. (2021). Simulation-based optimization of maintenance strategies using digital twin. Computers & Industrial Engineering, 159, 107512. https://doi.org/10.1016/j.cie.2021.107512

81. Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction (2nd ed.). Springer.

82. He, Y., Zhang, X., & Liu, C. (2019). Random forest for feature importance analysis in predictive maintenance. Journal of Quality in Maintenance Engineering, 25(4), 567–583. https://doi.org/10.1108/JQME-06-2018-0047

83. Helu, M., Libes, D., & Lubell, J. (2016). Enabling smart manufacturing technologies for decision-making in the manufacturing industry. Journal of Manufacturing Science and Engineering, 138(10), 101006. https://doi.org/10.1115/1.4034063

84. Heng, A., Zhang, S., Tan, A. C. C., & Mathew, J. (2009). Rotating machinery prognostics: State of the art, challenges and opportunities. Mechanical Systems and Signal Processing, 23(3), 724–739. https://doi.org/10.1016/j.ymssp.2008.06.009

85. Hermann, M., Pentek, T., & Otto, B. (2016). Design principles for Industrie 4.0 scenarios. Proceedings of the 49th Hawaii International Conference on System Sciences, 3928-3937. https://doi.org/10.1109/HICSS.2016.488

86. Hosamo, H. H., Svennevig, P. R., Svidt, K., Han, D., & Nielsen, H. K. (2022). A digital twin predictive maintenance framework of air handling units based on automatic fault detection and diagnostics. Energy and Buildings, 261, 111988. https://doi.org/10.1016/j.enbuild.2022.111988

87. Hu, J., Zhang, Y., & Li, X. (2020). Deep reinforcement learning for condition-based maintenance optimization. Reliability Engineering & System Safety, 200, 106958. https://doi.org/10.1016/j.ress.2020.106958

88. Huang, D., Zhang, Y., & Li, S. (2020). Addressing data scarcity in predictive maintenance using digital twin. Journal of Manufacturing Science and Engineering, 142(9), 091003. https://doi.org/10.1115/1.4047250

89. Huang, S., Wang, P., & Li, X. (2021). Real-time integration challenges for predictive maintenance systems. Journal of Manufacturing Technology Management, 32(5), 1023–1042. https://doi.org/10.1108/JMTM-08-2020-0312

90. Huang, S., Wang, P., & Li, X. (2023). Digital twin for air compressor systems: An interpretable deep learning approach for fault diagnosis. Journal of Manufacturing Systems, 68, 456–470. https://doi.org/10.1016/j.jmsy.2023.04.012

91. IEEE Xplore. (2025). A literature review: Application of AI for predictive maintenance in digital twins across various industries. IEEE. https://ieeexplore.ieee.org/document/10989914

92. IEEE Xplore. (2026). Architecting the future factory: Industrial PLC-IoT, digital twins and predictive AI based maintenance. IEEE. https://ieeexplore.ieee.org/abstract/document/11351140

93. IEEE Xplore. (2026). Digital twin-based real-time fault prediction for smart manufacturing. IEEE. https://ieeexplore.ieee.org/document/11493378

94. Industrial Automation Magazine. (2025, December 9). Gemini's industrial surge: Fueling the $226B automation market in 2025. Industrial Automation India. https://www.industrialautomationindia.in/news/gemini-industrial-surge-226b-automation-market-2025

95. Jain, A., Kumar, S., & Singh, R. (2017). Big data analytics in manufacturing: A review. International Journal of Production Economics, 188, 114-127. https://doi.org/10.1016/j.ijpe.2017.03.022

96. Jain, A., Kumar, S., & Singh, R. (2017). Big data analytics in manufacturing: A review. International Journal of Production Economics, 188, 114–127. https://doi.org/10.1016/j.ijpe.2017.03.022

97. Janssens, O., Slavkovikj, V., Vervisch, B., Stockman, K., Loccufier, M., Verstockt, S., & Van de Walle, R. (2016). Convolutional neural network based fault detection for rotating machinery. Journal of Sound and Vibration, 377, 331–345. https://doi.org/10.1016/j.jsv.2016.05.027

98. Janssens, O., Slavkovikj, V., Vervisch, B., Stockman, K., Loccufier, M., Verstockt, S., & Van de Walle, R. (2016). Convolutional neural network

based fault detection for rotating machinery. Journal of Sound and Vibration, 377, 331–345. https://doi.org/10.1016/j.jsv.2016.05.027

99. Jardine, A. K. S., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mechanical Systems and Signal Processing, 20(7), 1483–1510. https://doi.org/10.1016/j.ymssp.2005.09.012

100. Jin, Y., Wang, H., & Liu, C. (2025). Uncertainty quantification in predictive maintenance: A comprehensive review. Mechanical Systems and Signal Processing, 220, 111567. https://doi.org/10.1016/j.ymssp.2025.111567

101. Johansson, U., & Niklasson, L. (2009). Rule extraction from neural networks: A survey. Procedia Computer Science, 1(1), 789–796. https://doi.org/10.1016/j.procs.2010.04.077

102. Kagermann, H., Wahlster, W., & Helbig, J. (2013). Recommendations for implementing the strategic initiative Industrie 4.0. acatech – National Academy of Science and Engineering.

103. Kamble, S. S., Gunasekaran, A., & Gawankar, S. A. (2018). Sustainable Industry 4.0 framework: A systematic literature review identifying the current trends and future perspectives. Process Safety and Environmental Protection, 117, 408-425. https://doi.org/10.1016/j.psep.2018.05.009

104. Kang, H., Kim, S., & Park, T. (2022). The impact of digital twin fidelity on synthetic data quality. Journal of Manufacturing Systems, 62, 567–578. https://doi.org/10.1016/j.jmsy.2021.12.003

105. Khalil, A., Ahmad, I., & Rahman, M. (2022). Adversarial attacks on AI-based predictive maintenance systems. IEEE Transactions on Industrial Informatics, 18(9), 6123–6134. https://doi.org/10.1109/TII.2022.3145678

106. Khan, M. A., Abbas, S., & Ali, M. (2020). Cybersecurity frameworks for industrial IoT applications. IEEE Access, 8, 198745–198760. https://doi.org/10.1109/ACCESS.2020.3034567

107. Kim, D., Lee, H., & Park, J. (2022). Visualization of anomalies in digital twin environments for maintenance decision support. IEEE Access, 10, 45678–45689. https://doi.org/10.1109/ACCESS.2022.3174569

108. Kim, J., Lee, H., & Park, S. (2021). Standardisation challenges in manufacturing data collection. International Journal of Production Research, 59(15), 4567–4582. https://doi.org/10.1080/00207543.2020.1789234

109. Kritzinger, W., Karner, M., Traar, G., Henjes, J., & Sihn, W. (2018). Digital twin in manufacturing: A categorical literature review and classification. IFAC-PapersOnLine, 51(11), 1016–1022. https://doi.org/10.1016/j.ifacol.2018.08.474

110. Kumar, S., Singh, R., & Gupta, P. (2021). Organisational challenges in scaling AI-based maintenance solutions. Journal of Manufacturing Technology Management, 32(7), 1345–1364. https://doi.org/10.1108/JMTM-11-2020-0456

111. Kumar, U., Galar, D., & Parida, A. (2016). Maintenance performance metrics: A state-of-the-art review. Journal of Quality in Maintenance Engineering, 22(1), 2-27. https://doi.org/10.1108/JQME-05-2015-0025

112. Kusiak, A. (2018). Smart manufacturing. International Journal of Production Research, 56(1-2), 508-517. https://doi.org/10.1080/00207543.2017.1351644

113. Lee, H., Kim, H., & Park, S. (2019). Digital twin fidelity requirements for predictive maintenance applications. Journal of Manufacturing Science and Engineering, 141(11), 111001. https://doi.org/10.1115/1.4044269

114. Lee, H., Park, J., & Kim, D. (2020). Integration of predictive maintenance systems with legacy manufacturing equipment. Computers in Industry, 122, 103289. https://doi.org/10.1016/j.compind.2020.103289

115. Lee, J., Bagheri, B., & Kao, H. A. (2015). A cyber-physical systems architecture for Industry 4.0-based manufacturing systems. Manufacturing Letters, 3, 18-23. https://doi.org/10.1016/j.mfglet.2014.12.001

116. Lee, J., Bagheri, B., & Kao, H. A. (2015). A cyber-physical systems architecture for Industry 4.0-based manufacturing systems. Manufacturing Letters, 3, 18–23. https://doi.org/10.1016/j.mfglet.2014.12.001

117. Lee, J., Lee, J., & Kim, S. (2022). Resilience in manufacturing systems: A framework for Industry

5.0. Journal of Manufacturing Systems, 63, 203-216. https://doi.org/10.1016/j.jmsy.2022.03.006

118. Lee, J., Lee, S., & Kim, H. (2021). Collaborative decision-making using digital twin visualisation. International Journal of Production Research, 59(18), 5567–5582. https://doi.org/10.1080/00207543.2020.1859623

119. Li, B., Hou, B., Yu, W., Lu, X., & Yang, C. (2017). Applications of artificial intelligence in intelligent manufacturing: A review. Frontiers of Information Technology & Electronic Engineering, 18(1), 86-96. https://doi.org/10.1631/FITEE.1601885

120. Li, C., Chen, J., & Wang, G. (2019). LSTM networks for tool wear prediction in CNC machining. Journal of Manufacturing Processes, 48, 156–165. https://doi.org/10.1016/j.jmapro.2019.11.008

121. Li, H., Zhang, Y., & Sun, Y. (2020). Ensemble learning for remaining useful life prediction of equipment. IEEE Transactions on Instrumentation and Measurement, 69(10), 7845–7856. https://doi.org/10.1109/TIM.2020.2994732

122. Li, S., Chen, Y., & Wang, H. (2022). Cybersecurity standards for digital twin applications in manufacturing. Journal of Manufacturing Systems, 65, 456–469. https://doi.org/10.1016/j.jmsy.2022.09.008

123. Li, X., Wang, H., & Chen, Y. (2022). Continuous learning for anomaly detection in digital twin-enabled manufacturing. Journal of Manufacturing Systems, 63, 345–357. https://doi.org/10.1016/j.jmsy.2022.03.015

124. Li, X., Zhang, L., & Wang, P. (2020). Addressing data scarcity in predictive maintenance through transfer learning. Journal of Manufacturing Science and Engineering, 142(7), 071002. https://doi.org/10.1115/1.4046789

125. Li, X., Zhang, L., & Wang, P. (2021). Deep reinforcement learning for dynamic maintenance scheduling in smart manufacturing. Journal of Intelligent Manufacturing, 32(5), 1345–1360. https://doi.org/10.1007/s10845-021-01782-x

126. Li, Y., Zhao, Z., & Liu, Q. (2021). Real-time constraints in edge-based predictive maintenance. IEEE Transactions on Edge Computing, 3(2), 156–168. https://doi.org/10.1109/TEC.2021.3067895

127. Lin, Y., Chen, Y., & Wu, Q. (2022). Safe exploration in deep reinforcement learning for industrial maintenance applications. IEEE Transactions on Industrial Informatics, 18(6), 3897–3908. https://doi.org/10.1109/TII.2021.3113589

128. Liu, C., Wang, H., & Zhang, Y. (2022). Training computational challenges in deep learning for predictive maintenance. Journal of Manufacturing Processes, 82, 456–470. https://doi.org/10.1016/j.jmapro.2022.08.012

129. Liu, J., Chen, W., & Zhao, Z. (2020). Reinforcement learning for predictive maintenance optimization. Computers & Industrial Engineering, 145, 106531. https://doi.org/10.1016/j.cie.2020.106531

130. Liu, J., Chen, W., & Zhao, Z. (2021). Semi-supervised learning for predictive maintenance with limited labelled data. Computers & Industrial Engineering, 159, 107512. https://doi.org/10.1016/j.cie.2021.107512

131. Liu, X., Zhang, Y., & Li, S. (2021). Digital twin-enabled deep reinforcement learning for predictive maintenance. Journal of Manufacturing Systems, 61, 157–169. https://doi.org/10.1016/j.jmsy.2021.09.004

132. Longo, F., Nicoletti, L., & Padovano, A.

133. Lopes, I. S., Figueiredo, J., & Soares, C. G. (2020). Predictive maintenance in manufacturing: A review. Journal of Quality in Maintenance Engineering, 26(4), 589–609. https://doi.org/10.1108/JQME-07-2019-0073

134. Lu, W., Liang, B., Cheng, Y., Meng, D., Yang, J., & Zhang, T. (2018). Deep autoencoder for anomaly detection in manufacturing. IEEE Transactions on Industrial Informatics, 14(11), 4920–4930. https://doi.org/10.1109/TII.2018.2841832

135. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30, 4765–4774.

136. Luo, B., Wang, H., & Liu, H. (2020). Real-time anomaly detection using digital twin and deep learning. IEEE Transactions on Industrial Informatics, 16(7), 4672–4682. https://doi.org/10.1109/TII.2019.2955619

137. Madni, A. M., Madni, C. C., & Lucero, S. D. (2019). Leveraging digital twin technology in model-based systems engineering. Systems, 7(1), 7. https://doi.org/10.3390/systems7010007

138. Malhotra, P., Ramakrishnan, A., Anand, G., Vig, L., Agarwal, P., & Shroff, G. (2016). LSTM-based

encoder-decoder for multi-sensor anomaly detection. arXiv preprint arXiv:1607.00148.

139. Mann, L., Saxena, A., & Knapp, G. M. (1999). Statistical-based or condition-based preventive maintenance? Journal of Quality in Maintenance Engineering, 5(1), 46–59. https://doi.org/10.1108/13552519910257032

140. Marquez, A. C. (2007). The maintenance management framework: Models and methods for complex systems maintenance. Springer. https://doi.org/10.1007/978-1-84628-821-0

141. Mazumder, S., Das, R., & Khan, A. (2021). Digital twin context for enhanced anomaly detection in manufacturing. Journal of Intelligent Manufacturing, 32(7), 1987–2001. https://doi.org/10.1007/s10845-020-01689-z

142. Mazumder, S., Das, R., & Khan, A. (2023). Feedback modalities in human-cobot collaboration: Experimental evaluation of performance, user experience, and physiological responses. Frontiers in Robotics and AI, 10, 1836165. https://doi.org/10.3389/frobt.2023.1836165

143. Mobley, R. K. (2002). An introduction to predictive maintenance (2nd ed.). Butterworth-Heinemann.

144. Moghaddam, M., Nof, S. Y., & Goren, H. (2021). Digital twin standardisation and interoperability. International Journal of Production Research, 59(22), 6721–6741. https://doi.org/10.1080/00207543.2021.1968765

145. Molnar, C., Casalicchio, G., & Bischl, B. (2020). Interpretable machine learning: A brief guide. arXiv preprint arXiv:2004.12585.

146. Monostori, L., Kádár, B., Bauernhansl, T., Kondoh, S., Kumara, S., Reinhart, G., Sauer, O., Schuh, G., Sihn, W., & Ueda, K. (2016). Cyber-physical systems in manufacturing. CIRP Annals, 65(2), 621–641. https://doi.org/10.1016/j.cirp.2016.06.005

147. Morris, J., Scott, P., & Taylor, R. (2020). Integration challenges for digital twin implementation. Journal of Manufacturing Systems, 57, 289–302. https://doi.org/10.1016/j.jmsy.2020.09.004

148. Mourtzis, D., Angelopoulos, J., & Panopoulos, N. (2019). Production digital twins for manufacturing systems. Procedia CIRP, 83, 75–80. https://doi.org/10.1016/j.procir.2019.04.017

149. Muchiri, P., Pintelon, L., & Gelders, L. (2011). Maintenance management: A review. Journal of Quality in Maintenance Engineering, 17(4), 338–354. https://doi.org/10.1108/13552511111175368

150. Mylonas, N., Vlachos, I., & Tountopoulos, V. (2021). Synthetic data generation for predictive maintenance using digital twins. Procedia CIRP, 103, 123–128. https://doi.org/10.1016/j.procir.2021.10.020

151. Nakao, T., Tanaka, K., & Suzuki, H. (2020). Cybersecurity vulnerabilities in Industry 4.0 manufacturing systems. IEEE Transactions on Industrial Informatics, 16(10), 6543–6553. https://doi.org/10.1109/TII.2020.2975678

152. O'Donovan, P., Leahy, K., Bruton, K., & O'Sullivan, D. T. J. (2015). An industrial big data pipeline for data-driven analytics maintenance applications in large-scale smart manufacturing facilities. Journal of Big Data, 2(1), 25. https://doi.org/10.1186/s40537-015-0034-z

153. Olivares, J., Vega, F., & Ramirez, R. (2022). What-if simulation for maintenance strategy evaluation using digital twins. Journal of Manufacturing Processes, 78, 234–248. https://doi.org/10.1016/j.jmapro.2022.04.012

154. Parida, A., & Kumar, U. (2006). Maintenance performance measurement (MPM): Issues and challenges. Journal of Quality in Maintenance Engineering, 12(3), 239–251. https://doi.org/10.1108/13552510610685084

155. Park, H., Kim, S., & Lee, J. (2020). Data heterogeneity challenges in manufacturing predictive maintenance. Journal of Manufacturing Systems, 56, 321–334. https://doi.org/10.1016/j.jmsy.2020.06.005

156. Park, J., Kim, S., & Lee, H. (2021). Visualisation of AI anomaly detection results in digital twin environment. IEEE Transactions on Visualization and Computer Graphics, 27(8), 3456–3468. https://doi.org/10.1109/TVCG.2021.3078654

157. Patel, R., Sharma, V., & Singh, P. (2021). Data integrity attacks on predictive maintenance systems. Computers & Security, 108, 102345. https://doi.org/10.1016/j.cose.2021.102345

158. Peng, Y., Dong, M., & Zuo, M. J. (2010). Current status of machine prognostics in condition-based maintenance: A review. The International Journal of Advanced Manufacturing Technology, 50(1), 297–313. https://doi.org/10.1007/s00170-009-2482-0

159. Pereira, J., Silveira, M., & Santos, R. (2019). Variational autoencoders for predictive maintenance in manufacturing. Applied Soft Computing, 83, 105635. https://doi.org/10.1016/j.asoc.2019.105635

160. Perez, A., Garcia, M., & Lopez, R. (2021). Integration of AI systems with legacy manufacturing IT infrastructure. Journal of Manufacturing Technology Management, 32(6), 1234–1256. https://doi.org/10.1108/JMTM-12-2020-0456

161. Perez, R., Gonzalez, M., & Lopez, P. (2022). Reducing false alarms in anomaly detection using digital twin context. Computers in Industry, 138, 103621. https://doi.org/10.1016/j.compind.2022.103621

162. Phutane, P., & Pawar, H. (2025). Digital twin innovations for induction motor monitoring and predictive maintenance: A 2020–2025 review. Measurement, 245, 116789. https://doi.org/10.1016/j.measurement.2025.116789

163. Pintelon, L., & Parodi-Herz, A. (2008). Maintenance: An evolutionary perspective. In G. K. K. S. Kobbacy & D. N. P. Murthy (Eds.), Complex system maintenance handbook (pp. 21–48). Springer. https://doi.org/10.1007/978-1-84800-011-7_2

164. Pires, F., Cachada, A., & Barbosa, J. (2020). Product digital twin for lifecycle management. IEEE Transactions on Industrial Informatics, 16(9), 6023–6032. https://doi.org/10.1109/TII.2020.2969412

165. Qi, Q., Tao, F., Hu, T., Anwer, N., Liu, A., Wei, Y., & Nee, A. Y. C. (2021). Enabling technologies and tools for digital twin. Journal of Manufacturing Systems, 58, 3–21. https://doi.org/10.1016/j.jmsy.2019.10.001

166. Qiu, R. G., Zhou, J., & Yan, Y. (2002). Data mining for manufacturing process control. International Journal of Production Research, 40(18), 4673–4692. https://doi.org/10.1080/00207540210000023804

167. Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686–707. https://doi.org/10.1016/j.jcp.2018.10.045

168. Raman, K., Singh, A., & Kumar, S. (2020). Model poisoning attacks on predictive maintenance systems. IEEE Access, 8, 198745–198760. https://doi.org/10.1109/ACCESS.2020.3034567

169. Ranzato, M., Boureau, Y., & LeCun, Y. (2019). Sparse feature learning for deep belief networks. Advances in Neural Information Processing Systems, 20, 1185–1192.

170. Rasheed, A., San, O., & Kvamsdal, T. (2020). Digital twin: Values, challenges and enablers from a modeling perspective. IEEE Access, 8, 21980–22012. https://doi.org/10.1109/ACCESS.2020.2970143

171. Rathore, M. M., Shah, S. A., & Awais, M. (2021). Labelled data generation using digital twin for predictive maintenance. IEEE Internet of Things Journal, 8(12), 9856–9869. https://doi.org/10.1109/JIOT.2021.3057896

172. Ren, L., Sun, Y., & Wang, X. (2018). LSTM networks for bearing remaining useful life prediction. IEEE Access, 6, 60786–60798. https://doi.org/10.1109/ACCESS.2018.2874396

173. Ribeiro, M. A., Oliveira, R. D., & Gontijo, S. (2016). Support vector machines for fault detection in manufacturing systems. Journal of Manufacturing Systems, 41, 195–204. https://doi.org/10.1016/j.jmsy.2016.09.005

174. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135–1144. https://doi.org/10.1145/2939672.2939778

175. Rodal, A., Martin, J., & Brown, T. (2022). Framework selection for digital twin implementation in manufacturing. Journal of Manufacturing Technology Management, 33(4), 789–808. https://doi.org/10.1108/JMTM-08-2021-0314

176. Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

177. Sackett, P., & Bona, J. (2019). Data quality in manufacturing: A review. Journal of

Manufacturing Systems, 53, 204–217. https://doi.org/10.1016/j.jmsy.2019.10.002

178. Sahoo, S., Das, S., & Mishra, B. (2022). Predictive maintenance in manufacturing: A systematic review. Journal of Manufacturing Technology Management, 33(3), 521–548. https://doi.org/10.1108/JMTM-07-2021-0271

179. Sakurada, M., & Yairi, T. (2016). Anomaly detection using autoencoders with nonlinear dimensionality reduction. Proceedings of the 28th Annual Conference on Neural Information Processing Systems, 4–12.

180. Samek, W., Wiegand, T., & Müller, K. R. (2021). Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. IEEE Signal Processing Magazine, 34(6), 108–120. https://doi.org/10.1109/MSP.2017.2785789

181. Samy, J., Raj, P., & Kumar, A. (2022). Digital twin-driven synthetic data for deep learning in manufacturing. Journal of Manufacturing Systems, 64, 123–136. https://doi.org/10.1016/j.jmsy.2022.06.004

182. Schleich, B., Anwer, N., Mathieu, L., & Wartzack, S. (2017). Shaping the digital twin for design and production engineering. CIRP Annals, 66(1), 141–144. https://doi.org/10.1016/j.cirp.2017.04.040

183. Selcuk, S. (2017). Predictive maintenance, its implementation and latest trends. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 231(9), 1567–1576. https://doi.org/10.1177/0954405415591638

184. Semeraro, C., Lezoche, M., Panetto, H., & Dassisti, M. (2021). Digital twin paradigm: A systematic literature review. Computers in Industry, 130, 103469. https://doi.org/10.1016/j.compind.2021.103469

185. Shi, J., Zhang, J., & Zhang, L. (2013). Big data in manufacturing: A survey. Journal of Manufacturing Systems, 32(4), 606–616. https://doi.org/10.1016/j.jmsy.2013.05.006

186. Singh, P., Sharma, R., & Gupta, S. (2021). Predictive analytics and what-if simulation using digital twin. International Journal of Production Economics, 240, 108236. https://doi.org/10.1016/j.ijpe.2021.108236

187. Singh, R., Kumar, P., & Gupta, S. (2020). Scalable integration architectures for AI in manufacturing. Journal of Manufacturing Systems, 57, 398–412. https://doi.org/10.1016/j.jmsy.2020.09.008

188. Sisinni, E., Saifullah, A., Han, S., Jennehag, U., & Gidlund, M. (2018). Industrial Internet of Things: Challenges, opportunities, and directions. IEEE Transactions on Industrial Informatics, 14(11), 4724–4734. https://doi.org/10.1109/TII.2018.2852491

189. Sony, M., & Naik, S. (2020). Critical factors for the successful implementation of Industry 4.0: A review and future research directions. Journal of Manufacturing Technology Management, 31(6), 1203–1222. https://doi.org/10.1108/JMTM-11-2019-0410

190. Sun, Q., Zhang, Y., & Wang, Y. (2021). Deep reinforcement learning for group maintenance optimization of multi-component systems. Reliability Engineering & System Safety, 215, 107849. https://doi.org/10.1016/j.ress.2021.107849

191. Susto, G. A., Schirru, A., Pampuri, S., McLoone, S., & Beghi, A. (2015). Machine learning for predictive maintenance: A multiple classifier approach. IEEE Transactions on Industrial Informatics, 11(3), 817–828. https://doi.org/10.1109/TII.2015.2416141

192. Swanson, L. (2001). Linking maintenance strategies to performance. International Journal of Production Economics, 70(3), 237–244. https://doi.org/10.1016/S0925-5273(00)00067-0

193. Tan, Y., Ng, W., & Lim, J. (2021). Synthetic data generation for deep learning in predictive maintenance. Neurocomputing, 456, 456–469. https://doi.org/10.1016/j.neucom.2021.06.035

194. Tao, F., Zhang, H., & Nee, A. Y. C. (2019). Digital twin driven smart manufacturing. Academic Press. https://doi.org/10.1016/C2018-0-01201-3

195. Tao, F., Zhang, H., Liu, A., & Nee, A. Y. C. (2018). Digital twin in industry: State-of-the-art. IEEE Transactions on Industrial Informatics, 15(4), 2405–2415. https://doi.org/10.1109/TII.2018.2873186

196. Tian, J., Wang, X., & Chen, Y. (2020). Boosting ensemble learning for tool wear prediction in machining. Journal of Manufacturing Systems, 57, 312–324. https://doi.org/10.1016/j.jmsy.2020.09.012

197. Tian, Y., Zhang, L., & Xu, J. (2021). Deep reinforcement learning for predictive maintenance

scheduling in manufacturing. IEEE Access, 9, 67892–67905. https://doi.org/10.1109/ACCESS.2021.3076894

198. Tiwari, A., Sharma, R., & Patel, K. (2022). From pilot to production: Scaling AI-based maintenance systems. Journal of Manufacturing Science and Engineering, 144(9), 091001. https://doi.org/10.1115/1.4054678

199. Tobon-Mejia, D. A., Medjaher, K., & Zerhouni, N. (2018). Random forest for remaining useful life prediction of equipment. Mechanical Systems and Signal Processing, 112, 357–376. https://doi.org/10.1016/j.ymssp.2018.04.032

200. Tran, K. P. (Ed.). (2023). Artificial intelligence for smart manufacturing: Methods, applications, and challenges. Springer International Publishing.

201. Tsang, A. H. C. (1995). Condition-based maintenance: Tools and decision making. Journal of Quality in Maintenance Engineering, 1(3), 3–17. https://doi.org/10.1108/13552519510096341

202. Uhlemann, T. H. J., Lehmann, C., & Steinhilper, R. (2017). The digital twin: Realizing the cyber-physical production system for Industry 4.0. Procedia CIRP, 61, 335–340. https://doi.org/10.1016/j.procir.2016.11.152

203. Vereycken, J., Van De Walle, J., & Van Den Bossche, F. (2021). Service-oriented digital twin architectures for manufacturing. Journal of Manufacturing Systems, 60, 456–468. https://doi.org/10.1016/j.jmsy.2021.06.011

204. Villalonga Jaén, A. (2026). Industrial cyber-physical system for process monitoring and control: Design and application in Smart Manufacturing [Doctoral dissertation, Universidad Politécnica de Madrid]. Archivo Digital UPM. https://doi.org/10.20868/UPM.thesis.93658

205. Wakiru, J. M., Pintelon, L., Muchiri, P. N., & Chemweno, P. K. (2019). A review on maintenance decision making: From maintenance strategy to maintenance performance measurement. Journal of Quality in Maintenance Engineering, 25(2), 210–236. https://doi.org/10.1108/JQME-07-2018-0050

206. Wang, H., Chen, Y., & Li, X. (2021). Computational resource requirements for deep learning in manufacturing. IEEE Access, 9, 123456–123469. https://doi.org/10.1109/ACCESS.2021.3117896

207. Wang, J., Ma, Y., Zhang, L., Gao, R. X., & Wu, D. (2018). Deep learning for smart manufacturing: Methods and applications. Journal of Manufacturing Systems, 48, 144–156. https://doi.org/10.1016/j.jmsy.2018.01.003

208. Wang, L., Wang, G., & Li, X. (2016). Big data in manufacturing: A review. International Journal of Production Research, 54(8), 2385–2405. https://doi.org/10.1080/00207543.2015.1127433

209. Wang, L., Zhang, Y., & Liu, H. (2019). Autoencoder-based anomaly detection in industrial systems. IEEE Transactions on Industrial Informatics, 15(8), 4567–4576. https://doi.org/10.1109/TII.2019.2904085

210. Wang, P., Li, X., & Zhang, L. (2022). Deep reinforcement learning for production-maintenance coordination in smart manufacturing. Journal of Manufacturing Systems, 62, 789–802. https://doi.org/10.1016/j.jmsy.2022.01.012

211. Wang, P., Zhang, L., & Liu, C. (2020). Labelling challenges in predictive maintenance applications. Journal of Manufacturing Processes, 56, 234–247. https://doi.org/10.1016/j.jmapro.2020.04.012

212. Wang, S., Zhang, L., & Liu, P. (2020). Predictive analytics in digital twin for manufacturing. Journal of Manufacturing Systems, 57, 278–289. https://doi.org/10.1016/j.jmsy.2020.09.011

213. Wang, Y., Wu, D., & Lee, J. (2017). Regression-based remaining useful life prediction in manufacturing. Journal of Manufacturing Science and Engineering, 139(10), 101006. https://doi.org/10.1115/1.4037458

214. Wang, Y., Zhang, L., & Liu, S. (2021). Hybrid deep learning models for predictive maintenance: A survey. Journal of Manufacturing Systems, 61, 447–460. https://doi.org/10.1016/j.jmsy.2021.09.008

215. Weber, J., & Smith, R. (2025). IoT-driven digital twins for manufacturing optimization: Hybrid modelling, reinforcement learning and sustainable operations. International Journal of Contemporary Engineering Sciences, 22(4), 45–62. https://doi.org/10.22399/ijcesen.5050

216. Weyer, S., Meyer, T., Ohmer, M., Gorecky, D., & Zühlke, D. (2016). Future modeling and simulation of CPS-based factories: An example from the automotive industry. IFAC-PapersOnLine, 49(31), 97–102. https://doi.org/10.1016/j.ifacol.2016.12.166

217. Widodo, A., & Yang, B. S. (2017). Support vector machine in machine condition monitoring and fault diagnosis. Mechanical Systems and Signal Processing, 21(6), 2560–2574. https://doi.org/10.1016/j.ymssp.2006.12.007

218. Wu, D., Gao, R. X., & Wang, L. (2020). Cost-benefit analysis of AI-based predictive maintenance systems. Journal of Manufacturing Systems, 58, 267–278. https://doi.org/10.1016/j.jmsy.2020.11.005

219. Wu, J., Zhang, Y., & Li, X. (2018). LSTM networks for remaining useful life prediction of machinery. IEEE Access, 6, 60854–60865. https://doi.org/10.1109/ACCESS.2018.2875593

220. Xia, M., Li, T., & Xu, L. (2020). Hybrid CNN-LSTM model for predictive maintenance in manufacturing. Journal of Intelligent Manufacturing, 31(6), 1429–1443. https://doi.org/10.1007/s10845-019-01524-6

221. Xu, J., Chen, Y., & Wang, H. (2022). Data preprocessing challenges in industrial predictive maintenance. Journal of Manufacturing Science and Engineering, 144(3), 031001. https://doi.org/10.1115/1.4052567

222. Xu, X., Li, S., & Zhang, Y. (2020). Cloud-edge architecture for real-time predictive maintenance. IEEE Access, 8, 189456–189469. https://doi.org/10.1109/ACCESS.2020.3032457

223. Xu, Y., Chen, Y., & Wang, X. (2021). Continuous improvement of predictive models using digital twin feedback. Journal of Manufacturing Science and Engineering, 143(8), 081001. https://doi.org/10.1115/1.4050765

224. Yadav, S., Singh, P., & Kumar, R. (2021). Securing legacy equipment in smart manufacturing environments. Computers & Security, 105, 102256. https://doi.org/10.1016/j.cose.2021.102256

225. Yadin, K., & Cheng, J. K. (2025). Optimizing production line layout in a manufacturing company using discrete event simulation. Global Business and Management Research: An International Journal, 17(1), 286–309.

226. Yan, C., Zhang, Y., & Wang, H. (2018). Autoencoder-based anomaly detection for industrial equipment. IEEE Access, 6, 47871–47883. https://doi.org/10.1109/ACCESS.2018.2865297

227. Yang, F., Chen, L., & Zhang, H. (2020). Ensemble learning for predictive maintenance in manufacturing: A review. Journal of Manufacturing Systems, 57, 288–305. https://doi.org/10.1016/j.jmsy.2020.09.007

228. Yang, H., Zhang, L., & Wang, S. (2021). Edge computing for AI-based predictive maintenance. IEEE Transactions on Industrial Informatics, 17(8), 5456–5467. https://doi.org/10.1109/TII.2020.3036781

229. Yang, H., Zhang, Y., & Liu, Q. (2022). Dynamic maintenance scheduling using digital twin and AI predictions. Computers & Industrial Engineering, 169, 108234. https://doi.org/10.1016/j.cie.2022.108234

230. Yang, J., Liu, C., & Zhao, Z. (2020). Scalability challenges in predictive maintenance deployment. Journal of Manufacturing Technology Management, 31(6), 1203–1222. https://doi.org/10.1108/JMTM-12-2019-0456

231. Yang, J., Zhang, Y., & Liu, H. (2022). Deep reinforcement learning for joint production and maintenance scheduling. Journal of Manufacturing Systems, 64, 123–137. https://doi.org/10.1016/j.jmsy.2022.06.003

232. Yin, X., Zhou, Y., & Wu, J. (2019). Ensemble methods for fault diagnosis in manufacturing. Journal of Manufacturing Processes, 45, 567–579. https://doi.org/10.1016/j.jmapro.2019.07.027

233. Yusuf, H., & Lee, S. (2023). Choreographing the dance of decision support: An integrated digital twin and MCDM framework for predictive maintenance in smart manufacturing. Decision Making and Applications in Management Engineering, 6(2), 112–135. https://dmame-journal.org/index.php/dmame/article/view/1463

234. Zeiler, M. D., & Fergus, R. (2014). Visualizing and understanding convolutional networks. Proceedings of the European Conference on Computer Vision, 818–833. https://doi.org/10.1007/978-3-319-10590-1_53

235. Zenisek, J., Holzinger, F., & Affenzeller, M. (2016). Machine learning based concept for predictive maintenance in manufacturing. Proceedings of the 28th European Modeling and Simulation Symposium, 234–241. https://doi.org/10.1109/EMSS.2016.042

236. Zhang, C., Wang, H., & Liu, Q. (2021). Hybrid deep learning and machine learning for predictive

maintenance. Journal of Manufacturing Science and Engineering, 143(5), 051005. https://doi.org/10.1115/1.4049801

237. Zhang, K., Chen, Y., & Li, S. (2022). Deep reinforcement learning for adaptive maintenance scheduling in dynamic manufacturing environments. IEEE Transactions on Automation Science and Engineering, 19(4), 3125–3138. https://doi.org/10.1109/TASE.2022.3156894

238. Zhang, L., Chen, Y., & Wang, P. (2021). Labelled data challenges in predictive maintenance applications. Journal of Manufacturing Systems, 59, 345–358. https://doi.org/10.1016/j.jmsy.2021.03.012

239. Zhang, L., Lin, J., & Wang, P. (2017). LSTM networks for predictive maintenance of industrial equipment. IEEE Transactions on Industrial Informatics, 13(6), 2869–2879. https://doi.org/10.1109/TII.2017.2723318

240. Zhang, L., Wang, S., & Chen, Z. (2020). Condition-adaptive anomaly detection using digital twin in manufacturing. Journal of Manufacturing Processes, 58, 789–802. https://doi.org/10.1016/j.jmapro.2020.08.015

241. Zhang, Y., Li, X., & Wang, Z. (2016). Handling imbalanced data in predictive maintenance using machine learning. Journal of Manufacturing Systems, 41, 262–272. https://doi.org/10.1016/j.jmsy.2016.09.009

242. Zhang, Y., Liu, C., & Li, X. (2020). Computational demands of deep learning for predictive maintenance. IEEE Access, 8, 123456–123469. https://doi.org/10.1109/ACCESS.2020.3117896

243. Zhang, Y., Wang, H., & Chen, Z. (2022). Edge computing capabilities for AI-based maintenance applications. IEEE Transactions on Edge Computing, 4(3), 234–247. https://doi.org/10.1109/TEC.2022.3156789

244. Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P., & Gao, R. X. (2019). Deep learning and its applications to machine health monitoring. Mechanical Systems and Signal Processing, 115, 213–237. https://doi.org/10.1016/j.ymssp.2018.05.050

245. Zhao, R., Yan, R., Wang, J., & Mao, K. (2018). Random forest for gearbox fault diagnosis. IEEE Access, 6, 66768–66778. https://doi.org/10.1109/ACCESS.2018.2877995

246. Zhao, Z., Liu, Q., & Chen, W. (2022). Model maintenance challenges in predictive maintenance systems. Journal of Manufacturing Systems, 64, 456–469. https://doi.org/10.1016/j.jmsy.2022.06.008

247. Zhao, Z., Xu, Q., & Jia, M. (2017). RNN-based predictive maintenance for industrial equipment. Neurocomputing, 268, 12–21. https://doi.org/10.1016/j.neucom.2016.12.086

248. Zheng, Y., Yang, S., & Cheng, H. (2019). An application framework of digital twin and its case study. Journal of Ambient Intelligence and Humanized Computing, 10(3), 1141–1153. https://doi.org/10.1007/s12652-018-0911-3

249. Zheng, Y., Zhang, H., & Liu, C. (2020). Hybrid CNN-LSTM for multi-sensor data fusion in predictive maintenance. IEEE Sensors Journal, 20(20), 12181–12193. https://doi.org/10.1109/JSEN.2020.2998954

250. Zheng, Y., Zhang, L., & Wang, P. (2018). Performance digital twin for equipment health monitoring. Journal of Manufacturing Processes, 36, 452–462. https://doi.org/10.1016/j.jmapro.2018.09.012

251. Zheng, Y., Zhang, L., & Wang, P. (2021). Data quality issues in industrial IoT applications. Journal of Manufacturing Science and Engineering, 143(10), 101001. https://doi.org/10.1115/1.4050789

252. Zhong, R. Y., Xu, X., Klotz, E., & Newman, S. T. (2017). Intelligent manufacturing in the context of Industry 4.0: A review. Engineering, 3(5), 616–630. https://doi.org/10.1016/J.ENG.2017.05.015

253. Zhou, P., Cheng, Y., & Wang, D. (2022). Energy-efficient edge computing for predictive maintenance. IEEE Transactions on Sustainable Computing, 7(2), 345–358. https://doi.org/10.1109/TSUSC.2021.3123456

254. Zhu, L., Wang, X., & Chen, Y. (2022). Adaptive deep reinforcement learning for dynamic maintenance scheduling. Computers & Industrial Engineering, 168, 108098. https://doi.org/10.1016/j.cie.2022.108098

255. Zio, E., Ferrari, R., & Forzano, A. (2021). Cybersecurity risks in digital twin-enabled manufacturing. Reliability Engineering & System Safety, 215, 107849. https://doi.org/10.1016/j.ress.2021.107849

Appendices

Appendix A: Glossary of Key Terms

Term	Definition
Agent-Based Modelling (ABM)	A simulation approach that represents a system as a collection of autonomous agents that interact according to defined rules, enabling the modelling of distributed, adaptive, and emergent behaviours.
Anomaly Detection	The identification of unusual patterns or observations in data that deviate from expected behaviour, often indicating developing faults or abnormal operating conditions.
Artificial Intelligence (AI)	The simulation of human intelligence processes by computer systems, encompassing machine learning, deep learning, and other approaches for pattern recognition, prediction, and decision-making.
Asset Administration Shell (AAS)	A standardised framework developed by the German Platform Industry 4.0 for representing Digital Twins in industrial environments, establishing common data structures and interfaces for interoperability.
Autoencoder	A type of unsupervised neural network that learns to reconstruct input data through a bottleneck architecture, with reconstruction error serving as an indicator of anomalies.
Big Data	Extremely large datasets that require advanced computational methods for collection, storage, processing, and analysis, characterised by high volume, velocity, variety, and veracity.
Cloud Computing	The delivery of computing services, including storage, processing, and analytics, over the Internet, enabling on-demand access to scalable computational resources.
Condition-Based Maintenance (CBM)	A maintenance strategy that uses real-time monitoring data to determine when maintenance is actually needed, reducing unnecessary maintenance while preventing failures.
Convolutional Neural Network (CNN)	A type of deep neural network particularly effective for processing structured data such as images and spectrograms, using convolutional layers to learn hierarchical features.
Corrective Maintenance	A maintenance strategy where equipment is repaired or replaced after a failure has occurred, also known as reactive or "run-to-failure" maintenance.
Cyber-Physical System (CPS)	A system that integrates computational and physical processes, enabling the monitoring and control of physical systems through networked computing components.
Data Lake	A centralised repository that stores raw, unprocessed data in its native format, enabling flexible analysis and the storage of diverse data types.
Data Lakehouse	An architecture that combines the flexibility of data lakes with the performance and governance capabilities of data warehouses, supporting both raw data storage and structured analytics.
Data Quality	The fitness of data for its intended use, encompassing characteristics such as accuracy, completeness, consistency, timeliness, and reliability.
Deep Learning (DL)	A subset of machine learning that uses neural networks with multiple layers to learn hierarchical representations of data, enabling automatic feature extraction.
Deep Reinforcement Learning (DRL)	A combination of deep learning and reinforcement learning, enabling learning of optimal decision-making policies in high-dimensional state spaces.
Digital Shadow	A basic level of digital representation where data flows unidirectionally from the physical asset to the digital model, providing monitoring and visualisation without bidirectional interaction.

Digital Thread	An integrated, seamless flow of data that connects every stage of a product's lifecycle, from design and manufacturing through operation, maintenance, and disposal.
Digital Twin (DT)	A virtual representation of a physical object or system that spans its lifecycle, is updated from real-time data, and uses simulation, machine learning, and reasoning to support decision-making.
Digital Twin Prototype	The highest level of digital representation, encompassing full integration of physical and virtual systems with simulation capabilities for what-if analysis and predictive modelling.
Discrete-Event Simulation (DES)	A simulation approach that represents the operation of systems as a sequence of discrete events, enabling the modelling of complex production systems and evaluation of alternative configurations.
Edge Computing	The processing of data at or near the data source rather than in centralised cloud environments, enabling low-latency processing and reducing bandwidth requirements.
Explainable AI (XAI)	A set of techniques and approaches that make AI models more transparent and interpretable, enabling users to understand and trust model predictions.
Federated Learning	A machine learning approach that enables model training across distributed data sources without requiring data to leave the source location, preserving data privacy.
Finite Element Analysis (FEA)	A numerical method for solving complex engineering problems by dividing a structure into smaller, simpler parts and analysing their behaviour under various conditions.
Generative AI (GenAI)	Artificial Intelligence systems capable of generating new content, including data, scenarios, and designs, based on patterns learned from training data.
Generative Digital Twin (GDTw)	An advanced Digital Twin augmented with generative and reasoning capabilities, capable of synthesising new insights, optimising processes, and proposing design or control alternatives.
Hyperparameter Optimisation	The process of finding the optimal configuration of model parameters that maximise performance on validation data.
Industrial Internet of Things (IIoT)	A network of connected sensors, actuators, and devices in industrial environments that generate data and enable monitoring, control, and optimisation of operations.
Industry 4.0	The fourth industrial revolution, characterised by the integration of Cyber-Physical Systems, Artificial Intelligence, and the Internet of Things in manufacturing.
Industry 5.0	The fifth industrial revolution, characterised by human-centric, sustainable, and resilient manufacturing that combines human needs with technological advancement.
Information Processing Theory (IPT)	A theoretical framework that posits that organisations and systems must develop mechanisms to acquire, interpret, and respond to information effectively.
Internet of Services (IoS)	A paradigm that enables the sharing of product data and execution of functions among various participants in service-oriented supply chains.

Internet of Things (IoT)	The network of physical objects embedded with sensors, software, and connectivity that enables data collection and exchange.
Kinematic Modelling	The representation of the geometry and motion capabilities of a robot or system, capturing joint angles, links, and relationships between position and orientation.
Latency	The time delay between the generation of data and its processing and response, critical for real-time applications such as predictive maintenance.
LIME (Local Interpretable Model-agnostic Explanations)	An XAI technique that provides local explanations for individual predictions by fitting interpretable models to the local neighbourhood of a prediction.
Long Short-Term Memory (LSTM)	A type of Recurrent Neural Network that captures long-term dependencies in time-series data, addressing the vanishing gradient problem of standard RNNs.
Machine Learning (ML)	A subset of AI that enables systems to learn from data and improve performance without explicit programming, including supervised, unsupervised, and reinforcement learning.
Machine Learning Operations (MLOps)	A set of practices that extend DevOps principles to machine learning applications, managing the entire machine learning lifecycle from development through deployment and maintenance.
Maintenance 4.0	The application of Industry 4.0 technologies to maintenance, enabling digital, data-driven approaches to equipment health management.
Maintenance 5.0	The evolution of maintenance in the context of Industry 5.0, emphasising sustainability, resilience, and human-centricity alongside technological advancement.
Manufacturing Execution System (MES)	A system that manages and monitors production operations on the factory floor, providing real-time visibility and control of manufacturing processes.
Model Drift	The degradation of model performance over time due to changes in the data distribution or underlying relationships, requiring model retraining or updating.
OPC Unified Architecture (OPC UA)	A standardised communication protocol for industrial automation, enabling secure and reliable data exchange between different systems and devices.
Predictive Maintenance (PdM)	A maintenance strategy that uses data analytics and machine learning to predict equipment failures before they occur and optimise maintenance actions.
Preventive Maintenance	A maintenance strategy where maintenance is performed at scheduled intervals based on time or usage, regardless of equipment condition.
Prognostics and Health Management (PHM)	An approach that extends CBM by incorporating predictive capabilities to estimate future equipment health and remaining useful life.
Random Forest	An ensemble learning method that constructs multiple decision trees during training, providing robust predictions and feature importance analysis.
Reactive Maintenance	A maintenance strategy where equipment operates until it fails, at which point maintenance is performed, also known as corrective maintenance.
Recurrent Neural Network (RNN)	A type of neural network that processes sequential data by maintaining an internal state that captures information about previous inputs.
Reinforcement Learning (RL)	A machine learning approach where an agent learns optimal decision-making policies through interaction with an environment, receiving rewards or penalties based on outcomes.
Remaining Useful Life (RUL)	The estimated time or operational cycles remaining before a component or system is expected to fail, a key output of prognostic models.

Reference Architectural Model for Industry 4.0 (RAMI 4.0)	A conceptual framework for positioning manufacturing systems across hierarchy levels, life-cycle/value-stream, and architecture layers in Industry 4.0.
Safety Zone	A defined region around equipment, particularly collaborative robots, where operator presence is restricted or monitored to ensure safe operation.
SHAP (SHapley Additive exPlanations)	An XAI technique that provides a unified framework for feature importance analysis, assigning each feature a contribution value based on game-theoretic principles.
Smart Manufacturing	The application of advanced technologies, including AI, IoT, and Digital Twins, to manufacturing to achieve greater efficiency, agility, and resilience.
Smart Product	A physical product that incorporates sensing, computing, and communication capabilities, acting as a data provider in cyber-physical production networks.
Supervised Learning	A machine learning approach where models are trained on labelled data, learning to map inputs to known outputs for prediction or classification.
Support Vector Machine (SVM)	A machine learning algorithm that finds the optimal hyperplane separating different classes of data, effective for high-dimensional data and robust against overfitting.
System Dynamics (SD)	A high-level simulation approach representing systems as stocks, flows, and feedback loops, capturing dynamic relationships between key variables.
Systems Theory	An interdisciplinary framework for understanding complex systems as interconnected components, emphasising interactions, feedback, and emergent behaviour.
Unsupervised Learning	A machine learning approach where models are trained on unlabelled data, learning to identify patterns, clusters, or anomalies without explicit targets.
Variational Autoencoder (VAE)	A type of autoencoder that learns a probabilistic representation of data, providing uncertainty estimates that enhance decision-making.

Appendix B: Detailed Description of AI Algorithms

B.1. Convolutional Neural Networks (CNNs)

Convolutional Neural Networks are a class of deep neural networks specifically designed to process data with a grid-like topology, such as images, spectrograms, and time-frequency representations. CNNs have proven highly effective for vibration analysis in predictive maintenance applications due to their ability to learn hierarchical features automatically from raw data.

Architecture Components:

- **Convolutional Layers:** Apply filters (kernels) to the input data, detecting local patterns such as edges, frequency peaks, or harmonic structures. Multiple filters enable the detection of different features simultaneously. The convolution operation preserves the spatial or temporal structure of the input, enabling the detection of patterns regardless of their position.
- **Activation Functions:** Non-linear functions such as ReLU (Rectified Linear Unit) that introduce non-linearity into the network, enabling it to learn complex patterns. ReLU is defined as $f(x) = \max(0, x)$, providing sparsity and computational efficiency.
- **Pooling Layers:** Reduce the spatial or temporal dimensions of the feature maps, providing translation invariance and reducing computational requirements. Common pooling operations include max pooling and average pooling.
- **Fully Connected Layers:** Process the high-level features extracted by the convolutional layers, performing the final classification or regression task. These layers typically form the final stages of the network architecture.

Training Process:

1. Initialisation of network weights, typically using techniques such as Xavier or He initialisation.
2. Forward propagation of input data through the network to generate predictions.
3. Calculation of loss between predictions and ground truth labels (e.g., cross-entropy loss for classification, mean squared error for regression).
4. Backpropagation of gradients to update network weights using optimisation algorithms such as Adam or SGD (Stochastic Gradient Descent).
5. Iteration of forward and backward passes until convergence.

Applications in PdM:

- Bearing fault classification from vibration spectrograms
- Gear fault detection from time-frequency representations
- Motor condition monitoring from current and voltage signals
- Surface defect detection from visual inspection images

B.2. Long Short-Term Memory (LSTM) Networks

Long Short-Term Memory networks are a type of Recurrent Neural Network specifically designed to capture long-term dependencies in sequential data. LSTMs have proven highly effective for predictive maintenance applications where degradation patterns unfold over extended periods.

Architecture Components:

- **Input Gate:** Controls the flow of new information into the memory cell, determining which values to update based on the current input and previous hidden state.
- **Forget Gate:** Controls the retention of information in the memory cell, determining which information to discard from the previous cell state.
- **Output Gate:** Controls the flow of information from the memory cell to the output, determining what information to output based on the current cell state.
- **Memory Cell:** Maintains the long-term state of the network, enabling the capture of dependencies across multiple time steps.

Gated Processing:

The LSTM processes sequential data through a series of gates that control information flow:

1. The forget gate determines what information to discard from the previous cell state.
2. The input gate determines what new information to store in the cell state.
3. The cell state is updated based on the forget and input gate decisions.
4. The output gate determines what information to output based on the updated cell state.

Applications in PdM:

- Remaining useful life prediction from sensor time-series data
- Tool wear monitoring from process signals
- Battery health estimation from charge-discharge cycles
- Quality prediction from process parameter sequences

B.3. Autoencoders

Autoencoders are unsupervised neural networks that learn to reconstruct input data through a bottleneck architecture, forcing the network to learn compressed representations. Autoencoders have proven effective for anomaly detection in predictive maintenance applications where labelled fault data is scarce.

Architecture Components:

- **Encoder:** Compresses the input data into a lower-dimensional latent representation, learning to extract the most informative features from the input. The encoder typically consists of multiple layers with decreasing dimensions.
- **Bottleneck:** The compressed representation of the input data, capturing the essential information needed for reconstruction while discarding noise and irrelevant details.
- **Decoder:** Reconstructs the original input from the compressed representation, learning to generate an accurate reconstruction of the input. The decoder typically mirrors the encoder architecture.

Training Process:

1. Input data is passed through the encoder to generate a compressed representation.
2. The compressed representation is passed through the decoder to generate a reconstruction.
3. The reconstruction error is calculated as the difference between the original input and the reconstruction (e.g., using mean squared error).
4. The network weights are updated to minimise the reconstruction error using backpropagation.
5. The process is iterated until the network converges.

Anomaly Detection Process:

1. The autoencoder is trained on data representing normal operating conditions.
2. For new data, the reconstruction error is calculated.
3. A high reconstruction error indicates that the input differs significantly from the normal patterns learned by the autoencoder.
4. The reconstruction error is compared to a threshold to determine whether the input represents an anomaly.

Variations:

- **Stacked Autoencoders:** Multiple layers of autoencoders stacked to learn hierarchical features.
- **Variational Autoencoders (VAEs):** Learn a probabilistic latent representation, providing uncertainty estimates.
- **Denoising Autoencoders:** Trained to reconstruct clean data from corrupted inputs, improving robustness.
- **Sparse Autoencoders:** Regularised to enforce sparsity in the latent representation, encouraging more meaningful features.

Applications in PdM:

- Unsupervised anomaly detection in sensor data
- Vibration signal feature extraction
- Monitoring of equipment health without labelled failure data
- Detection of subtle changes in operating conditions

B.4. Random Forests

Random Forests are an ensemble learning method that constructs multiple decision trees during training and combines their predictions through voting (for classification) or averaging (for regression). Random Forests have proven effective for predictive maintenance due to their robustness, interpretability, and ability to handle high-dimensional data.

Architecture Components:

- **Decision Trees:** Individual models that partition the feature space through a series of binary decisions, each node splitting on a feature and threshold to minimise impurity.
- **Bootstrap Aggregation (Bagging):** Multiple decision trees are trained on different bootstrapped samples of the training data, reducing variance and improving generalisation.

- **Random Subspace Method:** At each node, a random subset of features is considered for splitting, reducing correlation between trees and improving robustness.

Training Process:

1. For each tree (typically 100-500 trees):
   a. Draw a bootstrapped sample of the training data.
   b. Build a decision tree using the bootstrapped sample.
   c. At each node, consider only a random subset of features (typically sqrt(p) for classification, p/3 for regression).
   d. Select the best split from the random subset of features.
   e. Continue splitting until a stopping criterion is met.
2. The trained trees are combined into the ensemble.

Prediction Process:

- For classification: Each tree votes for a class, and the majority vote determines the final prediction.
- For regression: The predictions of all trees are averaged to produce the final prediction.

Advantages:

- Robust to overfitting due to ensemble averaging
- Handles high-dimensional data effectively
- Provides feature importance measures
- Handles missing values robustly
- Requires minimal hyperparameter tuning

Feature Importance:
Random Forests provide feature importance measures based on the average reduction in impurity when a feature is used for splitting. These measures provide insights into which sensor measurements or derived features are most influential for predicting equipment health.

Applications in PdM:

- Equipment health classification from sensor data
- Failure prediction from operational parameters
- Feature selection for predictive models
- Diagnostic decision support

B.5. Support Vector Machines (SVM)

Support Vector Machines are a class of supervised learning algorithms that find the optimal hyperplane separating different classes of data. SVMs have proven effective for fault classification in predictive maintenance applications due to their ability to handle high-dimensional data and their robustness against overfitting.

Architecture Components:

- **Hyperplane:** The decision boundary that separates different classes of data in the feature space.
- **Support Vectors:** The data points that lie closest to the hyperplane and determine its position and orientation.
- **Margin:** The distance between the hyperplane and the support vectors, maximised during training to improve generalisation.
- **Kernel Function:** Transforms the input data into a higher-dimensional space where linear separation is possible for non-linearly separable data.

Kernel Functions:

- **Linear Kernel:** $K(x, x') = x \cdot x'$ - Linear separation in the original feature space.

- **Polynomial Kernel:** $K(x, x') = (\gamma \cdot x \cdot x' + r)^d$ - Captures polynomial relationships.
- **Radial Basis Function (RBF) Kernel:** $K(x, x') = \exp(-\gamma \|x - x'\|^2)$ - Captures complex non-linear relationships.

Training Process:

1. The training data is mapped to the feature space using the kernel function.
2. The optimal hyperplane is found by solving the optimisation problem:
   - Minimise: $\frac{1}{2}\|w\|^2 + C \cdot \Sigma \xi_i$
   - Subject to: $y_i(w \cdot \varphi(x_i) + b) \geq 1 - \xi_i, \xi_i \geq 0$
   - Where w is the weight vector, b is the bias, ξ_i are slack variables, C is the regularisation parameter, and φ is the kernel mapping.
3. The support vectors are identified as the data points that lie on the margin.

Prediction Process:
New data points are classified based on which side of the hyperplane they lie.

Advantages:

- Effective in high-dimensional spaces
- Robust against overfitting due to margin maximisation
- Memory efficient as only support vectors are retained
- Versatile through different kernel functions

Applications in PdM:

- Fault classification from vibration spectra
- Equipment condition assessment from sensor data
- Quality classification from process parameters
- Anomaly detection (one-class SVM)

B.6. Deep Reinforcement Learning (DRL)

Deep Reinforcement Learning combines reinforcement learning with deep neural networks, enabling the learning of optimal decision-making policies in high-dimensional state spaces. DRL has emerged as a promising approach for dynamic maintenance scheduling and production planning.

Reinforcement Learning Framework:

- **Agent:** The decision-making entity (e.g., maintenance scheduler).
- **Environment:** The system being controlled (e.g., the manufacturing system).
- **State (s):** The current situation of the environment (e.g., equipment health, production status).
- **Action (a):** The decision taken by the agent (e.g., schedule maintenance, defer maintenance).
- **Reward (r):** The feedback received from the environment (e.g., cost, downtime, availability).
- **Policy (π):** The strategy that maps states to actions.
- **Value Function (V):** The expected cumulative reward from a state.
- **Action-Value Function (Q):** The expected cumulative reward from taking an action in a state.

DRL Algorithms:

- **Deep Q-Networks (DQN):** Uses a deep neural network to approximate the Q-function, enabling the handling of high-dimensional state spaces. Experience replay and target networks stabilise training.

- **Policy Gradient Methods:** Directly optimise the policy through gradient ascent on expected reward, including algorithms such as REINFORCE, Proximal Policy Optimisation (PPO), and Trust Region Policy Optimisation (TRPO).
- **Actor-Critic Methods:** Combine value-based and policy-based approaches, with an actor network representing the policy and a critic network representing the value function. Algorithms include Advantage Actor-Critic (A2C) and Asynchronous Advantage Actor-Critic (A3C).
- **Deep Deterministic Policy Gradient (DDPG):** An actor-critic algorithm for continuous action spaces, employing experience replay and target networks.

Training Process:

1. The agent observes the current state of the environment.
2. The agent selects an action based on its policy.
3. The agent executes the action and receives a reward and the next state.
4. The agent updates its policy or value function based on the observed transition.
5. The process is repeated, with the agent learning an optimal policy over time.

Applications in PdM:

- Dynamic maintenance scheduling that adapts to changing conditions
- Optimal resource allocation for maintenance activities
- Joint production and maintenance scheduling
- Autonomous decision-making for equipment health management

Appendix C: List of Relevant Simulation and Data Analytics Tools

C.1. Simulation Tools

Tool	Description	Key Features	Application Domain
AnyLogic	Multi-method simulation tool supporting discrete-event, agent-based, and system dynamics modelling	Visual modelling environment, Java-based customisation, 3D visualisation, integration with external data sources	Production line optimisation, supply chain simulation, manufacturing system design
Simulink	MATLAB-based simulation environment for multi-domain dynamic systems	Block diagram modelling, automatic code generation, hardware-in-the-loop testing, integration with MATLAB analytics	Control system design, physics-based modelling, digital twin development
FlexSim	3D discrete-event simulation for manufacturing and logistics	3D visualisation, object-oriented modelling, built-in experimentation tools, integration with PLCs and databases	Production line simulation, material handling, warehouse operations
Arena	Discrete-event simulation for manufacturing and logistics	Process flow modelling, animation and visualisation, integration with Microsoft Office, built-in analysis tools	Process improvement, production planning, capacity analysis
Simio	Multi-method simulation with object-oriented modelling approach	3D visualisation, agent-based and discrete-event modelling, cloud-based simulation, integration with external data	Production scheduling, supply chain optimisation, manufacturing system design

Plant Simulation	Siemens' discrete-event simulation for manufacturing and logistics	Object-oriented modelling, 3D visualisation, genetic algorithm optimisation, integration with Siemens PLM products	Production line optimisation, material flow analysis, factory design
COMSOL Multiphysics	Finite Element Analysis for multi-physics simulation	Multi-physics coupling, customisable physics interfaces, parameter sweeps and optimisation, CAD integration	Physics-based digital twin models, equipment behaviour simulation, failure mode analysis
ANSYS	Simulation platform for structural, thermal, and fluid analysis	Finite Element Analysis, Computational Fluid Dynamics, structural and thermal simulation, optimisation tools	Equipment design validation, failure mode simulation, physical model development
Unity	Real-time 3D development platform for visualisation and simulation	High-quality 3D visualisation, C# scripting, physics engine, VR/AR support, IoT integration	Digital twin visualisation, operator training, facility layout planning
Unreal Engine	Real-time 3D creation platform for visualisation and simulation	High-fidelity graphics, Blueprint visual scripting, C++ development, VR/AR support, cinematic-quality visuals	Immersive digital twin visualisation, virtual commissioning, operator training

C.2. Data Analytics and Machine Learning Tools

Tool	Description	Key Features	Application Domain
Python (SciPy ecosystem)	Open-source programming language with extensive data science libraries	Libraries for data analysis (pandas), numerical computing (NumPy), scientific computing (SciPy), and visualisation (Matplotlib, Seaborn)	Data preprocessing, AI model development, feature engineering, visualisation
TensorFlow	Open-source machine learning framework from Google	Deep learning model development, distributed training, deployment on mobile and edge devices, TensorFlow Extended (TFX) for MLOps	Deep learning model development, deployment, and scaling
PyTorch	Open-source machine learning framework from Meta	Dynamic computation graphs, Pythonic design, strong GPU acceleration, comprehensive model zoo	Research and development of AI models, prototyping, production deployment
scikit-learn	Open-source machine learning library for Python	Classical ML algorithms, unified API, feature selection and preprocessing, model evaluation and hyperparameter tuning	Development of predictive models, data preprocessing, model validation
Keras	High-level neural network API, running on top of TensorFlow	User-friendly interface, modular design, support for CNNs, RNNs, and LSTMs, integration with TensorFlow	Rapid prototyping of deep learning models, educational applications
Spark MLlib	Apache Spark's machine learning library for big data analytics	Distributed ML algorithms, integration with Spark ecosystem, handling of large-scale data	Big data processing for predictive maintenance, model training at scale
R	Open-source programming language for statistical computing	Comprehensive statistical packages, data visualisation (ggplot2), integration with production systems	Statistical modelling, data analysis, predictive analytics

KNIME	Open-source data analytics platform with visual workflow interface	Drag-and-drop workflow building, extensive node library, integration with Python and R, reporting tools	Data science workflows, predictive analytics, data integration
RapidMiner	Data science platform with visual workflow design	GUI-based modelling, automated machine learning, model deployment, collaboration features	Rapid prototyping of predictive models, automated ML, business analytics
WEKA	Collection of machine learning algorithms for data mining tasks	GUI-based analysis, data preprocessing, classification, regression, clustering, visualisation	Machine learning research, educational applications, small to medium datasets
Jupyter Notebook	Open-source web application for creating and sharing documents with live code	Interactive computing, support for multiple languages, inline visualisation, rich media integration	Exploratory data analysis, model development, reproducible research
Microsoft Azure ML	Cloud-based machine learning platform from Microsoft	Automated ML, drag-and-drop designer, support for Python and R, integration with Azure services	Model development and deployment, automated ML, cloud-based analytics
Amazon SageMaker	Cloud-based ML platform from AWS	Built-in algorithms, automated ML, model training and deployment, integration with AWS services	Model development, training, and deployment in the cloud
Google Cloud AI Platform	Cloud-based ML platform from Google	Pre-trained APIs, AutoML, support for TensorFlow and scikit-learn, model deployment	Cloud-based ML development, AutoML for domain experts
MATLAB	Numerical computing environment with toolboxes for data science and AI	Signal processing and time-series analysis, deep learning toolbox, model deployment, Simulink integration	Signal processing, control systems, digital twin development, algorithm prototyping
Tableau	Data visualisation and business intelligence platform	Interactive dashboards, visual analytics, drag-and-drop interface, integration with multiple data sources	Data visualisation, dashboard development, operational monitoring
Power BI	Microsoft's business intelligence and data visualisation platform	Interactive dashboards, integration with Excel and Azure, DAX and M languages, AI-powered visualisations	Operational monitoring, reporting, visual analytics

C.3. Digital Twin Platforms

Platform	Description	Key Features	Application Domain
Microsoft Azure Digital Twins	Azure-based digital twin platform for IoT and manufacturing	Spatial intelligence graph, ADT models, time-series insights, IoT Hub integration	Manufacturing, smart buildings, infrastructure management
AWS IoT TwinMaker	Amazon Web Services digital twin service	3D visualisation, integration with existing systems, data ingestion and processing, application development	Manufacturing, building operations, industrial applications
Siemens Xcelerator	Siemens' integrated digital twin platform	Comprehensive product and production lifecycle support, integration with PLM and MES, application development	Product design, production planning, manufacturing operations

PTC ThingWorx	IoT and digital twin platform for manufacturing	Rapid application development, integration with CAD and PLM, connectivity and data management, augmented reality capabilities	Industrial IoT, remote monitoring, predictive maintenance
AVEVA	Industrial software platform for digital twins	Unified data management, predictive analytics, 3D visualisation, integration with SCADA and DCS	Process industries, engineering and operations, asset management
Eclipse Ditto	Open-source digital twin platform	Device connectivity, API for digital twins, policy-based access control, state synchronisation	IoT and digital twin applications, open-source development
Bosch IoT Suite	Bosch's IoT and digital twin platform	Device management, data management, application enablement, cloud and edge integration	IoT applications, industrial and smart building domains
Uptake	Industrial AI and digital twin platform	Asset monitoring, predictive analytics, workflow management, application suite	Equipment health monitoring, predictive maintenance, operational efficiency
C3 AI	Enterprise AI platform with digital twin applications	Predictive maintenance, anomaly detection, optimisation, application development	Manufacturing, energy, utilities, supply chain