PEOPLE’S DEMOCRATIC REPUBLIC OF ALGERIA
الجمهــــوريـــــة الجـزائــريـــــة الديــمـقـراطيـــــة الشـعـبيـــــة
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH
وزارة التعليــم العالـــي والـبحــث العلمــي
HIGHER SCHOOL OF COMPUTER SCIENCE AND DIGITAL TECHNOLOGIES-BEJAIA
المــــدرسة العـــــليا في علــــوم وتكنــولوجيــات الاعــــلام الالـــي بجـــاية

Dissertation Submitted to the Department Of Computer Science in Partial Fulfillment of the Requirements for Master’s Degree in Computer Science

Specialty: Artificial Intelligence and Data Sciences

Submitted By:

BELAYALI Rezkia **OURARI Tinhinane**

Predictive maintenance: state of the art based on Machine learning methods

Supervised by:

Dr. DAOUDI Meroua ESTIN

Members of jury:

- **Dr.** MEDJOUDJ Rafik President ESTIN
- **Dr.** SEBAA Abderrazak Examiner ESTIN
- **Mrs.** MELCHANE Selestine Examiner ESTIN
- **Mrs.** AZZOUGUER Dalila Examiner ESTIN

Academic year: 2023/2024

Abstract:

This study explores the evolving role of machine learning in revolutionizing predictive maintenance (PdM) within Industry 4.0, emphasizing the transition from traditional methods to advanced, data-driven approaches, particularly highlighting deep learning's transformative impact. It examines key technologies such as IoT (Internet of Things) sensors for real-time vibration analysis and addresses the efficacy of data-driven models, stressing the importance of managing data quality. The study also explores state-of-the-art approaches that integrate both single-model and multi-model frameworks, combining machine learning (ML) with physics-based models and statistical techniques. This integrated approach enhances anomaly detection, fault classification, and estimation of remaining useful life (RUL), contributing to a robust PdM framework designed for Industry 4.0 environments.

Keywords: Predictive maintenance, industry 4.0, machine learning, single model, multi-model, data-driven, IoT, vibration, RUL, anomaly detection, classification.

Résumé :

La maintenance prédictive (PdM) a évolué à travers les révolutions industrielles, passant de stratégies réactives et préventives à des approches prédictives avancées. Cette vue d'ensemble examine les technologies de PdM telles que l'analyse des vibrations, l'analyse d'huile et l'imagerie thermique, et classe les modèles en catégories basées sur les connaissances, basées sur les données et basées sur la physique. L'apprentissage machine, en particulier l’apprentissage profond, joue un rôle crucial dans l'amélioration de la PdM. Le rapport aborde les défis de la qualité des données, y compris les données bruyantes et déséquilibrées, et explore l'intégration de l'IoT et des capteurs pour la collecte et l'analyse de données en temps réel. Les méthodes de PdM de pointe basées sur l'apprentissage machine sont examinées, mettant en lumière les avantages de la combinaison de différents modèles pour améliorer les résultats de maintenance et soulignant le potentiel transformateur de l'apprentissage machine dans diverses applications industrielles.

Mots clé : Maintenance prédictive, Industrie 4.0, apprentissage automatique, modèle simple, multi-modèle, basé sur les données, capteurs IoT, vibrations, RUL, détection des anomalies, classification.

Acknowledgements

First and foremost, we would like to express our deepest gratitude to Allah, the Almighty and the most Merciful, for granting us the strength, health, and perseverance to complete this work.

Secondly, we extend our sincere thanks to all those who have contributed, in one way or another, to the realization of this thesis.

In particular, we would like to warmly thank our supervisor, Dr. M. Daoudi, for her availability, her expert advice, and her invaluable guidance throughout our journey. Her encouragement and trust in us have been a great source of motivation.

We would also like to thank Phd student A. Mankouri for her invaluable help and advice. Her contribution has been of great importance to the progress of our work.

We are grateful to all those who have supported and accompanied us throughout this journey especially our school ESTIN.

Dedication:

To my dear mother, for your unconditional love, constant support, and sacrifices, thank you for being my source of inspiration and strength, for being strong in weakness, and for encouraging me to always strive for my best.

In memory of my father, who has left us, you remain forever in our hearts and thoughts. Your spirit continues to guide our steps every day.

To my two sisters, for your complicity, affection, and wisdom. You are much more than just sisters, you are pillars in my life, unwavering sources of support and constant comfort.

To my brother, for your friendship, which brightens even the darkest days. Your presence in my life is a constant source of joy and strength, and I am endlessly grateful for the bond we share.

Tinhinane OURARI

To Mom & Dad, my unwavering source of inspiration. Thank you for the countless sacrifices you've made to give me the opportunity to learn and grow. Your strength and love have guided me through every challenge.

To my siblings, my built-in cheerleaders. Thank you, Sister, for always supporting my decisions, no matter what. Thanks to my eldest brother for being my right hand, and to my youngest siblings for filling my days with laughter and joy.

To my friends, thank you for sharing life's journey with me. A special thanks to Asma for being a shoulder to lean on and a confidante who shares problems and worries.

Finally, to myself, thank you for not giving up. Thank you for working hard and persevering. I am proud of the person you have become.

Rezkia BELAYALI

Table of Contents

List of figures:

List of tables :

Acronyms

ANN: Artificial Neural Network
AUC-ROC: Area Under the Receiver Operating Characteristic Curve
AI: Artificial Intelligence
ARIMA: AutoRegressive Integrated Moving Average
BIM: Building Information Model
CART: Classification And Regression Trees
CNC: Computer Numerical Control
CPS: Cyber-Physical System
CSV: Comma-Separated Values
DBSCAN: Density-Based Spatial Clustering of Applications with Noise
DL: Deep Learning
DNN: Deep Neural Network
DT: Decision Tree
EM: Expectation-Maximization
FIR: Fourth Industrial Revolution
FM: Facility Management
GFMM: Gaussian Finite Mixture Model
GMM: Gaussian Mixture Model
Hz: Hertz
IIoT: Industry Internet of Things
KPI: Key Performance Indicators
LSTM: Long Short-term Memory
LOF: Local Outlier Factor
MAD: Mean Absolute Deviation
MATLAB: MATrix LABoratory
MC: Multiple Classifier
MCCV: Monte Carlo Cross Validation
MEP: Mechanical, Electrical, and Plumbing
ML: Machine Learning
MSE: Mean Squared Error
NMF: Non-negative Matrix Factorization
NASA: National Aeronautics and Space Administration
PCA: Principal Component Analysis
PC: Personal Computer
PdM: Predictive Maintenance
RBF: Radial Basis Function
RF: Random Forest
RMSE: Root Mean Square Error

RMS: Root Mean Square
RUL: Remaining Useful Life
SC: Supply Chains
SMOTE: Synthetic Minority Oversampling Technique
SOM: Self-Organizing Maps
SSL: Semi-Supervised Learning
SVR: Support Vector Regression
SVD: Singular Value Decomposition
SVM: Support Vector Machine
TCN: Temporal Convolutional Networks
WAV: Waveform Audio File Format
XDK: Cross Domain Development Kit

General Introduction:

The industrial landscape is undergoing a rapid transformation driven by automation, digitalization, and the relentless pursuit of efficiency. Industry 4.0, characterized by the integration of physical and digital worlds (Dalzochio et al., 2020), is at the forefront of this change. Within this dynamic environment, ensuring the reliable operation of industrial machinery remains paramount. From intricate manufacturing equipment to colossal processing systems, these machines are the backbone of global production, powering the creation of goods and services that fuel economic growth. While inherently reliable and high-performing, industrial machines are susceptible to wear, tear, and unforeseen failures. These breakdowns translate into significant challenges for industries striving to optimize their operations.

Traditionally, reactive or preventive maintenance strategies have dominated industrial settings. However, these approaches often lead to unplanned downtime, costly repairs, and under used assets (Jimenez et al., 2020). Thankfully, a paradigm shift is underway, with PdM emerging as a game-changer. PdM represents an innovative strategy that harnesses the power of advanced technologies, data analytics, and sensor networks to proactively predict equipment failures and schedule maintenance activities. This approach rests on the foundation of strategically placed sensors that continuously monitor machine health, performance, and condition, providing a wealth of data for informed decision-making (Adams et al., 2017; Jimenez et al., 2020). Industry 4.0 principles, such as interoperability and data-driven insights, are particularly well-suited for implementing advanced PdM techniques.

Despite the immense potential of PdM, its widespread implementation remains hindered by several challenges. Integrating sensor data with machine learning algorithms can be intricate. One key hurdle lies in effectively integrating vibration sensor data with machine learning algorithms. Vibration data, while rich in information about a machine's health, can be complex and noisy. **Can these challenges be overcome, and if so, what strategies can be employed to facilitate the wider adoption of PdM in industrial settings?**

This comparative study aims to explore the application of PdM in industrial settings within the framework of Industry 4.0 using machine learning models and sensors data. By analyzing existing literature, we will investigate various methodologies and approaches employed in the field. Our focus will be on:

1. Comparative Analysis of Industrial Maintenance Methods:

We will conduct a comprehensive review of existing industrial maintenance techniques used in Industry 4.0 environments. This analysis will compare and contrast the strengths, weaknesses, and applicability of different methods.

2. Sensor Technology for Continuous Monitoring:

We will meticulously examine the role of sensor technologies in Industry 4.0 for continuous monitoring of equipment health and performance. This includes exploring the integration of various sensors like vibration sensors, temperature sensors, and current sensors, and their contributions to real-time data acquisition for PdM applications.

3. Machine Learning for Predictive Analytics:

Building upon the foundation of Industry 4.0 principles, we will investigate the integration and effectiveness of machine learning algorithms in analyzing the vast amount of sensor data collected. We will explore how these algorithms can be trained to identify patterns and anomalies within the data that signal impending equipment failures, enabling proactive intervention and maintenance scheduling.

4. Bridging the Gap Between Theory and Practice:

This study will bridge the gap between theoretical advancements in PdM and practical implementation within industrial settings. By analyzing existing literature and drawing upon the insights from the previous objectives, we will develop practical recommendations for effective implementation of PdM strategies in Industry 4.0 environments.

This report is structured into three core chapters:

- **Chapter 1: Concepts and Definitions:** This chapter lays the groundwork by defining key terms like Industry 4.0, predictive maintenance. We will explore the various techniques used in PdM, providing a solid foundation for understanding the field.

- **Chapter 2: Machine Learning and Sensors:** This chapter delves deeper into the two pillars of PdM – machine learning algorithms and sensor technology. We will investigate the diverse range of machine learning algorithms employed in predictive maintenance and examine the critical role of sensor technology and Industrial Internet of Things (IIoT) frameworks in continuously monitoring equipment health for PdM applications.

- **Chapter 3: State-of-the-Art in Predictive Maintenance:** Building upon the knowledge established in the previous chapters, this chapter will discuss the various existing approaches in PdM. We will conduct a comparative analysis, examining the strengths, weaknesses, and applicability of different techniques, ranging from traditional methods like vibration analysis to cutting-edge deep learning architectures. Additionally, we will explore the concept of multi-model approaches that combine different techniques for even more robust and accurate prediction

Chapter 01: Concepts and definitions

I. Introduction:

This chapter explores the fundamental principles and definitions that form the foundation of predictive maintenance within the Industry 4.0 framework. It traces the evolution of industry, from its humble beginnings powered by water and steam to the sophisticated era of automation and robotics we find ourselves in today. Along this journey, various types of maintenance strategies have emerged, ranging from reactive responses to issues, to more proactive approaches such as preventive maintenance. However, it is PdM that stands out as a beacon of innovation in the modern industrial landscape. By harnessing the power of data analytics and advanced technology, PdM offers a forward-looking solution that not only addresses current maintenance needs but also anticipates and mitigates future issues, driving efficiency and cost savings across industries.

II. The industry revolution:

The Industrial Revolution was a period of major economic and social change that began in Great Britain in the late 18th century and spread throughout the world. It marked a shift from an agrarian, handicraft economy to one dominated by machine manufacturing and industrial production (Xu et al., 2018). This journey unfolds across four distinct eras, each characterized by remarkable advancements in technology and production methods. These epochs are visually depicted in Figure 1.

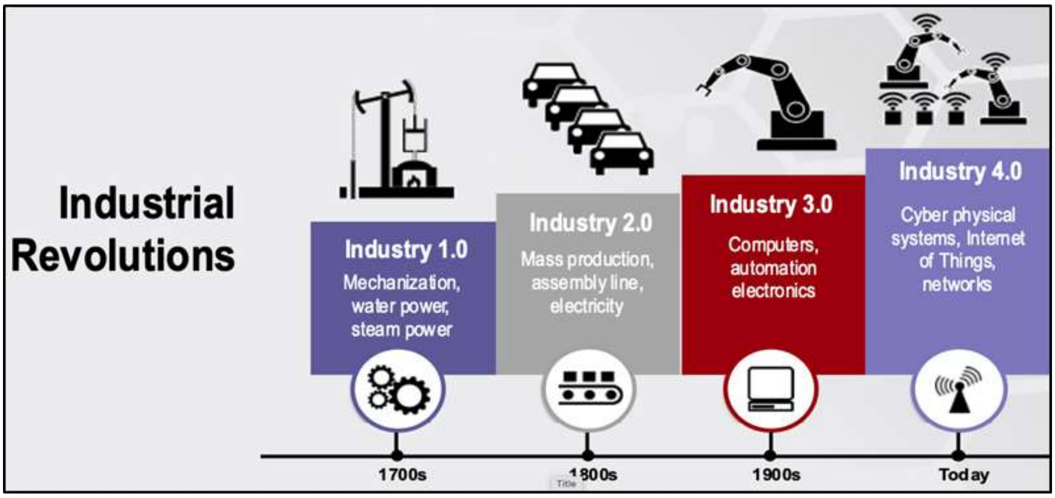

Figure 1 : : Industrial Revolution from 1.0 to 4.0 ((“Industrial Revolutions,” n.d.))

II.1 Industry 1.0: The Age of Steam (1760s)

The 18th century birthed the First Industrial Revolution. Steam power, a known technology, revolutionized industry by powering mechanized production. This leap, exemplified by the eightfold increase in textile output with steam-powered looms, dramatically boosted human productivity. Steam engines replaced muscle for weaving, paving the way for further advancements like steamships and locomotives. These innovations dramatically transformed transportation, enabling the movement of people and goods across vast distances in a fraction of the time (Poór, Ženíšek, et al., 2019).

II.2 Industry 2.0: The Rise of Electricity (1870s)

The 19th century ushered in the Second Industrial Revolution, marked by the rise of electricity and assembly line production. Inspired by the efficient disassembly process observed in a Chicago slaughterhouse (where pigs hung on conveyor belts and butchers each performed specific tasks), Henry Ford (1863-1947) revolutionized automobile manufacturing. He adopted the concept of mass production, breaking down car assembly into smaller steps on a conveyor belt. This innovation drastically increased production speed and reduced costs (Nabila et al., 2021).

II.3 Industry 3.0: The Digital Revolution (1970s)

The 1970s marked the dawn of the Third Industrial Revolution, characterized by the rise of partial automation. This revolution leveraged memory-programmable controls and computers to automate portions of the production process. Unlike earlier methods, these advancements paved the way for fully automated production lines, where tasks are completed by robots without human intervention. These robots operate based on programmed sequences, maximizing efficiency and minimizing human involvement (Poór, Ženíšek, et al., 2019).

II.4 Industry 4.0: cyber physical systems (2010s)

Also known as the Fourth Industrial Revolution (FIR), marks a recent revolution blurring the lines between the physical and digital worlds in manufacturing. This transformative era, driven by a convergence of technologies under development since the mid-1980s, integrates cutting-edge advancements like biotechnology, nanotechnology, artificial intelligence, and robotics with traditional processes (Poór, Ženíšek, et al., 2019). Coined in Germany in 2011, the term "Industry 4.0" emerged from a government initiative promoting connected manufacturing and digital convergence across industries, businesses, and other processes. It envisioned the "Internet of Things" (IoT) revolutionizing production organization by enabling a new interplay between humans and machines, and the application of digital technologies throughout manufacturing (Culot et al., 2020).

A key feature of Industry 4.0 is Cyber-Physical Systems (CPS). CPS excel in control, monitoring, transparency, and overall production efficiency. In simpler terms, CPS integrates computational power directly with physical machinery, enabling real-time data acquisition through the Internet of Things (IoT) infrastructure. This data is then analyzed to generate valuable insights and key performance indicators (KPIs) that empower real-time decision-making across supply chains (SC) (Morella et al., 2021).

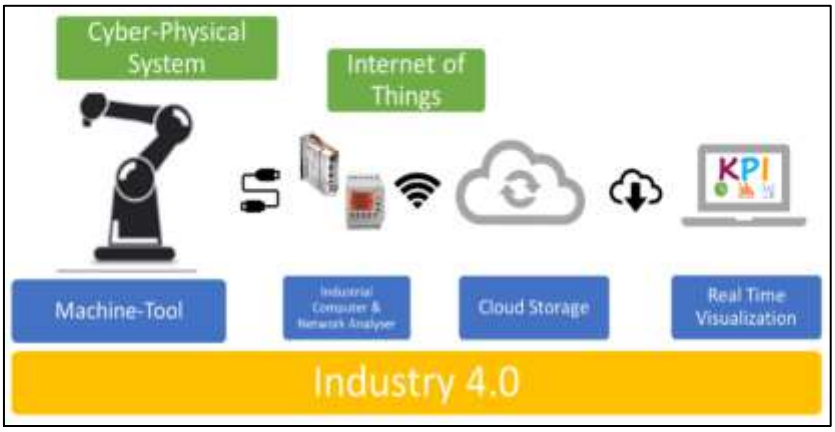

Figure 2 : Integration for real-time acquisition (Morella et al., 2021)

III. Industrial maintenance

III.1 Definition:

According to the Swedish standard SS-EN 13306, maintenance is described as: A set of technical, administrative, and managerial activities conducted throughout an item's life cycle to keep it in, or return it to, a condition where it can perform its required functions (Poór, Basl, et al., 2019).

III.2 Types of industrial maintenance:

III.2.1 Reactive Maintenance:

The fundamental principle guiding this maintenance policy is to postpone action until the equipment malfunctions, at which stage repairs or the replacement of faulty components are initiated (Bengtsson, 2004). This approach proves effective when equipment downtime doesn't affect production efficiency or when replacement costs are negligible, though such situations are rare. Furthermore, overseeing a maintenance department amidst continual crisis interventions presents significant hurdles. In the event of production halts, unplanned interventions become imperative, necessitating access to a diverse array of tools and spare parts. Ultimately, this maintenance strategy is deemed less effective (Sullivan et al., 2010). Its subtypes:

- **Curative Maintenance:** It fixes equipment after a breakdown to restore functionality (Rachidi et al., 2013).
- **Palliative Maintenance:** It prolongs the life of aging equipment by addressing wear and tear without fully resolving underlying issues (Rachidi et al., 2013).

III.2.2 Preventive Maintenance:

AFNOR (FD X 60-000) describes it as: Maintenance carried out at specified intervals or according to set criteria, aimed at lowering the likelihood of asset failure or operational degradation (Bengtsson, 2004).

It involves scheduled inspections, servicing, and repairs performed on equipment or systems at predetermined intervals to prevent failures or malfunctions. It aims to reduce the likelihood of unexpected breakdowns and extend the lifespan of assets (Poór, Ženíšek, et al., 2019).

- **Systematic Maintenance :**

The underlying principle of this policy is to plan maintenance activities at set intervals, determined either by the equipment's usage hours or by a predetermined schedule. This method ensures that repairs or part replacements are conducted proactively, preempting any actual failures (Bengtsson, 2004). This strategy is effective when the equipment isn't constantly in use and when personnel possess the necessary expertise to determine the optimal timing for intervention. However, this approach has its drawbacks if maintenance is performed either prematurely or belatedly. In such cases, the equipment may be idled unnecessarily or parts may be replaced before they have reached the end of their useful life (Poór, Ženíšek, et al., 2019).

- **Conditional Maintenance :**

It involves monitoring the operation of the asset and significant operational parameters, followed by necessary actions. This approach prioritizes the equipment's condition, acknowledging the limitations of depending only on indicators such as operating time or cycle count .When the equipment's condition deteriorates beyond a certain threshold, the risk of shutdown becomes significant enough to warrant scheduled maintenance intervention

(Bengtsson, 2004). Conditional maintenance offers the advantage of promptly detecting failures as they arise, providing a buffer period before equipment breakdown occurs.

III.2.3 Predictive Maintenance:

It is defined by the European standard (NF EN 13306 X 60-319) as follows: *"Conditional maintenance executed based on extrapolated forecasts from analysis and evaluation of significant parameters of the asset's degradation."* (Bengtsson, 2004)

It is a proactive approach to equipment care that leverages real-time data and condition monitoring to predict potential failures before they occur. Instead of relying on traditional time-based schedules or simply waiting for breakdowns, PdM uses sensors and analytics to monitor equipment health, identify early signs of wear and tear, and schedule maintenance actions only when necessary (Aremu et al., 2018). In essence, PdM shifts from a reactive "fix-it-when-it-breaks" mindset to a proactive "prevent-the-break" (Bengtsson, 2004) approach, leading to significant improvements in overall plant performance, cost-effectiveness, and safety (Poór, Ženíšek, et al., 2019).

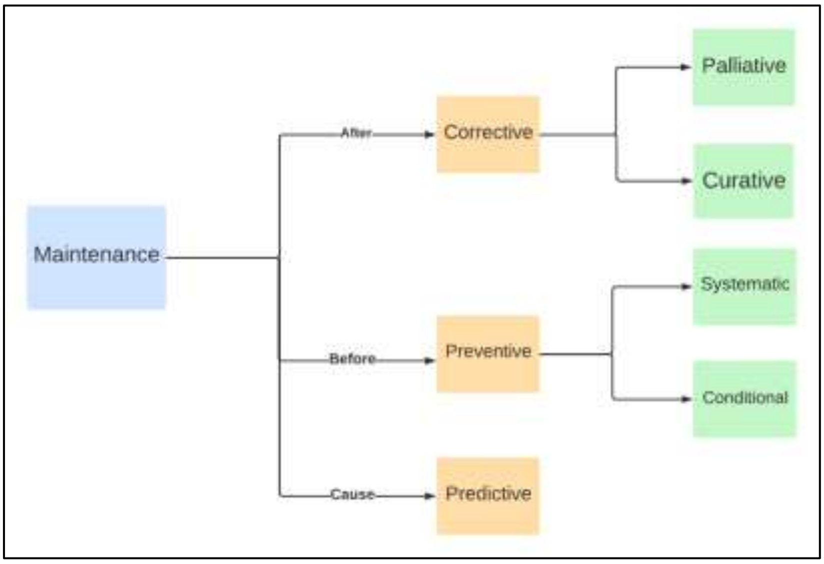

Figure 3 : Types of maintenance (strategies)

III.2.4 Comparison between the types of maintenance

Feature	**Reactive Maintenance**	**Preventive Maintenance**	**Predictive Maintenance**
Trigger	Equipment breakdown	Predetermined time/usage	Equipment state indicators suggest upcoming failure

Goal	Restore functionality after failure	Prevent failures through scheduled interventions	Prevent failures based on predicted wear and tear
Action	Repair or replace faulty components	Perform scheduled repairs/replacements	Perform maintenance based on predicted failure time
Effectiveness	Restores functionality, but doesn't prevent future failures	Effective in predictable scenarios, may be wasteful in others	Optimizes maintenance timing, minimizes downtime and costs
Cost	Varies depending on repair/replacement needed	Moderate, potential for unnecessary interventions	Highest due to data analysis and advanced technology
Suitability	Non-critical equipment, breakdown isn't costly	Simple equipment, stable usage patterns	High-value, complex equipment, critical operations
Advantages	Restores functionality quickly	Easy to implement, predictable costs	Minimizes downtime and costs, optimizes resource allocation
Disadvantages	Doesn't prevent future failures, reactive approach	Can be wasteful if overused, may not adapt to changing conditions	Complex implementation, requires advanced analytics

Table 1 : A comparison between the different types of maintenance

III.3 Predictive maintenance in today's dynamic environment:

Unplanned downtime due to equipment failure has severe consequences for industrial companies, impacting not just operations but also finances and reputation. Take **Amazon**'s infamous 49-minute outage in 2013, which resulted in a staggering $4 million loss in sales. This incident is just one example – studies show data center downtime incurs an average hourly cost of $138,000. Similarly, Operation and Maintenance (O&M) costs in industries like offshore wind and oil & gas can consume a significant portion of revenue, ranging from 20-35% and 15-70%, respectively (Sullivan et al., 2010). These figures highlight the immense financial burden associated with unplanned downtime across diverse industries. Implementing a well-defined and efficient maintenance strategy becomes crucial to address this issue. Here's where predictive maintenance comes in. allowing companies to:

- **Optimize equipment lifespan:** By identifying potential failures before they happen, companies can schedule maintenance proactively, maximizing equipment life.
- **Mitigate unplanned downtime:** Predictive maintenance helps avoid costly surprises by identifying and addressing issues before they cause disruptions.

- **Reduce energy consumption and costs:** By optimizing equipment performance, predictive maintenance can lead to significant energy savings.

Predictive maintenance is particularly valuable in complex manufacturing ecosystems with intricate interactions between different production activities. Its ability to anticipate issues becomes increasingly crucial as these ecosystems grow larger (Achouch et al., 2022; Poór, Basl, et al., 2019). As illustrated in Figure 4, the global adoption of this smart maintenance approach is expected to surge between 2022 and 2030.

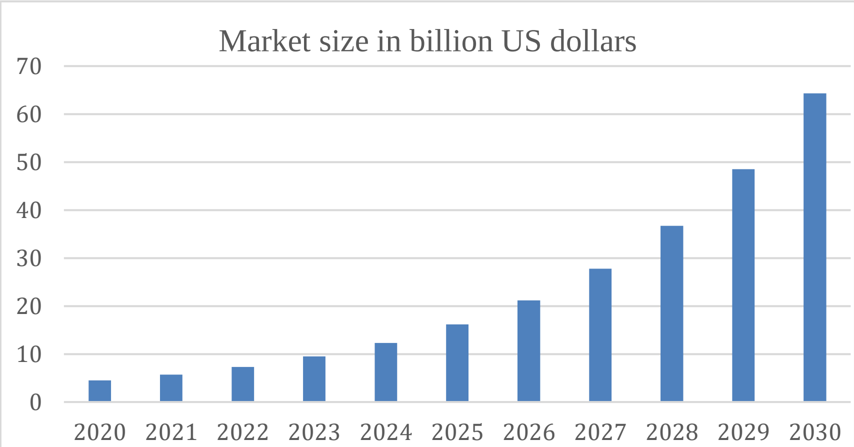

Figure 4 : Size of Maintenance 4.0 market worldwide in 2020 and 2021 with forecast for the future (Achouch et al., 2022)

III.4 A Brief History of Predictive Maintenance

The journey of predictive maintenance began around the turn of the 21st century, marking a significant shift from traditional, reactive maintenance practices. Initially, **organizations adopted a periodic or offline approach** to monitor asset condition. As highlighted in a 2001 study, "vibration measurements were taken periodically, typically once a month, with comparisons made against previous readings to identify potential issues." (Poór, Ženíšek, et al., 2019). This early stage of PdM was limited in its capabilities:

- **Infrequent monitoring:** Relying on monthly measurements meant potential problems could go undetected for extended periods, leading to unexpected failures.
- **Manual analysis:** Comparing data points manually was time-consuming and prone to human error.
- **Limited accessibility:** On-site data collection restricted access to real-time insights and remote monitoring capabilities.

However, the landscape of PdM has undergone a significant transformation. Today, organizations leverage a continuous or online approach for asset condition monitoring (Poór, Basl, et al., 2019). This involves:

- **Real-time data collection:** Sensors continuously gather data, providing a comprehensive and up-to-date picture of asset health.
- **Advanced analytics:** Sophisticated software analyzes data in real-time, identifying anomalies and predicting potential issues before they escalate.
- **Remote monitoring:** The Internet of Things (IoT) allows for connecting sensors to maintenance software, enabling remote monitoring and troubleshooting from any location.
- **Automated triggers:** When specific conditions are met, work orders for inspections and maintenance are automatically generated, ensuring prompt action can be taken.

III.5 Technologies of predictive maintenance:

PdM has become a cornerstone of effective asset management in various industries. This proactive approach uses diverse techniques to monitor equipment health, identify potential failures before they occur, and optimize maintenance schedules. Each PdM technique boasts distinct strengths and weaknesses, making a multifaceted approach crucial for achieving comprehensive asset monitoring (Scarf, 2007). Here are some technologies used

III.5.1 Vibration Analysis:

This widely employed technique uses sensors to capture subtle changes in equipment vibration patterns. These deviations can serve as early indicators of potential issues such as bearing wear, misalignment, or imbalances, allowing for preventative measures before critical failures arise (Hameed et al., 2009; Yan et al., 2014)

III.5.2 Oil Analysis:

By periodically analyzing oil samples from equipment, signs of potential problems like wear, contamination, or overheating can be detected. This early detection enables targeted maintenance actions, ultimately prolonging equipment lifespan and avoiding unplanned downtime (Hameed et al., 2009; Zhang et al., 1996).

III.5.3 Thermal Imaging:

This non-intrusive technique uses infrared cameras to identify areas of abnormal heat emission (hotspots) on equipment surfaces. These hotspots can signify potential issues like overheating electrical components, loose connections, or impending bearing failures, enabling targeted maintenance and preventing unforeseen equipment breakdowns (Hameed et al., 2009).

III.5.4 Motor Circuit Analysis:

This technique involves analyzing the electrical current and voltage characteristics of motors. Deviations from normal operating parameters can indicate potential problems like bearing wear, winding insulation breakdown, or impending motor failure. Early detection through motor circuit analysis allows for timely intervention and avoids costly downtime (Sharifi & Ebrahimi, 2011).

III.5.5 Acoustic Emission Testing:

This technique uses specialized sensors to detect high-frequency sounds emitted by equipment due to internal cracks, leaks, or loose components. By analyzing the characteristics of these high-frequency sounds, potential equipment issues can be identified before they escalate into catastrophic failures (Hameed et al., 2009; Kerkyras et al., n.d.).

III.5.6 Data-Driven PdM:

This overarching category encompasses techniques like Machine Learning Applications, Data Analytics, and Condition-Based Monitoring (CBM). All three methods leverage data from various sources, including sensor readings, historical records, and potentially external data like environmental factors. By using advanced analytics and machine learning algorithms, these techniques extract valuable insights from the data, enabling proactive maintenance strategies, optimized scheduling, and minimization of unnecessary interventions. These data-driven approaches offer a significant shift from traditional time-based maintenance, leading to improved equipment reliability, reduced downtime, and ultimately, enhanced operational efficiency and cost savings (Arcos Jiménez et al., 2017).

III.6 Types of predictive maintenance models:

Predictive maintenance has emerged as a crucial strategy for enhancing asset reliability and operational efficiency across industries. By harnessing advanced modeling techniques, the main types of models used for predictive maintenance are:

III.6.1 Single model:

III.6.1.1 Knowledge-based models:

Experience, captured through historical data and past maintenance actions, plays a crucial role. It allows the system to identify equipment faults, predict degradation patterns, and forecast potential component or system failures (Jimenez et al., 2020). It's divided into three main submodels: rule-based models, case-based models and fuzzy knowledge-based models.

- **Rule-based models:** Use IF-THEN rules to represent expert knowledge for diagnostic reasoning, root cause analysis, RUL estimation, as used for power circuit breakers (Hussain et al., 2015), fault diagnosis of wind turbines (Zhou et al., 2015), mining excavators (Kumar & Srivastava, 2012).
- **Case-based models:** Use stored cases of past problems/solutions to solve new problems through case retrieval and adaptation. Unlike rule-based reasoning, case-based reasoning can be used when the relations between facts cannot be declared explicitly (Vingerhoeds et al., 1995), as examples of applications of power equipment (Ma et al., 2015), induction motor fault diagnosis (Yang et al., 2004).
- **Fuzzy knowledge-based models:** Use fuzzy logic and linguistic rules to handle uncertainty in fault identification. Unlike Boolean logic, like a light switch, it is either on (true) or off (false). Fuzzy logic is like a dimmer switch. It allows for in-between states reflecting how we often perceive the world with varying degrees of truth (Vepa, 1992), as the fault diagnosis for grinding wheels (Baban et al., 2018).

III.6.1.2 Data-driven models:

The abundance of data generated by modern technical systems, coupled with advancements in computing power, has fueled the rise of data-driven models for maintenance purposes. These systems continuously measure and record a multitude of operational parameters, creating a vast dataset. This wealth of information can be harnessed explicitly or implicitly to extract valuable insights, including: Component Degradation Analysis, Real-time Health Assessment, and Remaining Useful Life. Data-driven models are classified in three groups: statistical models, stochastic models and machine learning models (Jimenez et al., 2020).

- **Statistical models:** Use statistical techniques like regression, Regression analysis: Use techniques like linear/exponential regression on sensor data for RUL estimation health assessment (Hu et al., 2012), lithium-ion batteries (Berecibar et al., 2016). ARMA models: Autoregressive models to forecast future values from time series data. Detection of gearbox deterioration (Zhan & Mechefske, 2007). Bayesian models: Use Bayesian theorem for parameter estimation, uncertainty management. Example direct RUL prediction (Mosallam et al., 2016)

- **Stochastic models:** Use stochastic processes like Gaussian processes: Non-linear regression using Gaussian processes. Example: fault diagnostics for wind turbines (Li et al., 2019). Markov chains: Model degradation as a Markov process for RUL computation. Example: semiconductor manufacturing (Kinghorst et al., 2017). Lévy processes: white LEDs (Yung et al., 2017).

- **Machine learning models:** is a branch of artificial intelligence that uses specialized learning algorithms to build models that extract hidden patterns from vast data, algorithms like SVMs, Neural networks (MLP, RNN, CNN)excel at fault diagnostics, RUL estimation through supervised learning as for nuclear power plant (Ayo-Imoru & Cilliers, 2018), bearings (Hinchi & Tkiouat, 2018). Unsupervised learning techniques like self-organizing maps (SOMs) offer additional insights by visualizing hidden relationships within complex data, particularly valuable for cyber-physical systems (Von Birgelen et al., 2018), Aircraft engine (Lu et al., 2019), SVR for RUL health assessment (Onel et al., 2018).

III.6.1.3 Physics-based models:

Use physics laws, finite element methods to model degradation phenomena like crack growth, fatigue. Example applications: rotor cages (Climente-Alarcon et al., 2017), robot health degradation (Qiao & Weiss, 2018), lithium-ion batteries (Downey et al., 2019).

III.6.2 Multi-model approaches:

Combine two or more of the above model types to overcome limitations of individual models. Fig. 5 presents a diagram of the potential combinations of multi-model approaches for predictive maintenance purposes (Jimenez et al., 2020). And here we go with some examples:

- **Knowledge-based & Data-driven:** combine Fuzzy logic and Markov chains for aero-engine prognostics (Liao, 2005)

- **Data-driven & Physics-based:** Physics-based model generates a health index. then analyze it by SVM, classify the health state and estimate the decline rate. Similarity analysis determines remaining useful life (Wang et al., 2012).
- **Data-driven & Physics-based:** uses a stochastic process (Wiener process) combined with a data analysis method (Principal Component Analysis) to model the deterioration of the

components that is fitted by an exponential physical degradation, and to estimate the remaining useful life (Le Son et al., 2013).

- **All three types combined**: Physics model + SVM + Fuzzy rules for rolling bearing diagnostics (Hong et al., 2009).

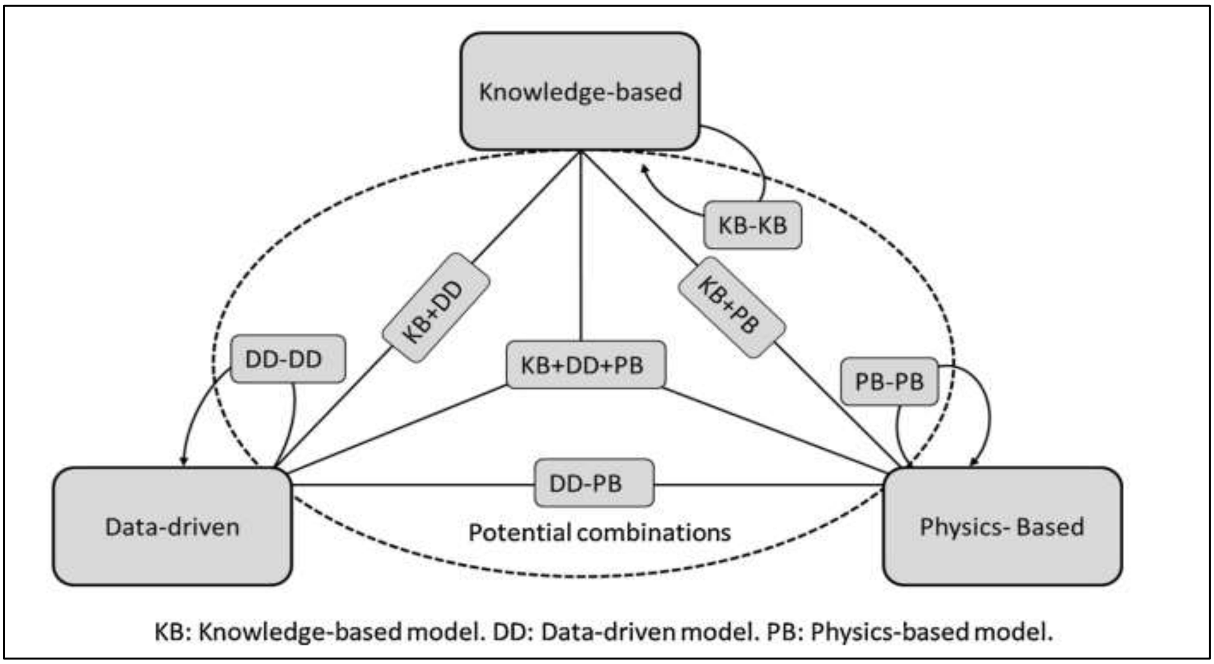

Figure 5 : Potential combinations for multimodel approaches (Jimenez et al., 2020)

IV. Conclusion

This chapter established the foundation for our exploration of predictive maintenance in Industry 4.0. We began by defining the concept of industrial revolutions, with a particular focus on Industry 4.0 and its emphasis on digitalization and automation. Next, we examined the various types of maintenance strategies, highlighting the shift from reactive to predictive approaches. We then explored the different forms of predictive maintenance starting from single-model approaches, followed by an in-depth discussion of composed approaches that leverage multiple models for enhanced accuracy.

The following chapter will build upon this knowledge by introducing the core algorithms used in machine learning and deep learning, the powerful tools that drive predictive maintenance. We will also explore the role of sensors in data acquisition, a crucial element for successful implementation.

Chapter 02: Machine learning and sensors

I. Introduction:

This chapter investigates the interconnected application of three key technologies within the domain of PdM. ML algorithms, encompassing a spectrum of complexity from basic models to sophisticated deep learning architectures, offer an analytical framework capable of extracting hidden insights from vast datasets. Sensor technology plays a crucial role in data acquisition, continuously capturing real-time parameters that delineate the health of equipment. Finally, the Industrial Internet of Things (IIoT) serves as the communication backbone, facilitating seamless data transmission and integration from these sensor networks.

II. Machine Learning:

II.1 Definition of Machine Learning:

Machine Learning is the field that involves developing algorithms and techniques to enable computers to acquire knowledge and improve their performance by analyzing and learning from data. Machine learning is a multifaceted concept with various definitions (Rachidi et al., 2013), but perhaps the most pertinent and widely accepted one comes from Tom M. Mitchell, a Professor of ML at the School of Computer Science, Carnegie Mellon University. Mitchell defines machine learning as "*A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.*" (Michalski et al., 2013). This definition is a concise way of explaining the essence of how machines can improve their performance over time through experience. It consists of:

- **Experience (E):** refers to the data or input that the machine is exposed to. It could be any form of information, ranging from structured data to unstructured text or images.
- **Task (T):** denotes the specific activity or problem that the machine is aiming to perform or solve. Tasks could vary widely, from recognizing handwriting to playing chess.
- **Performance Measure (P):** quantifies how well the machine is doing at the given task. It could be accuracy, precision, recall, F1-score… etc. Essentially, it's a metric used to evaluate the machine's performance.

II.2 The basic ML process:

The fundamental machine learning process consists of three key stages:

- **Data Input:** This stage involves using information from the past as historical data to help make decisions in the future (Lantz, 2019).
- **Abstraction:** The input data undergoes abstraction through an underlying algorithm, wherein it is represented in a more generalized and structured manner (Lantz, 2019).
- **Generalization:** The abstracted representation is generalized to create a framework for decision-making, enabling the model to make predictions or decisions on new, unseen data based on the patterns learned from the past data (Lantz, 2019).

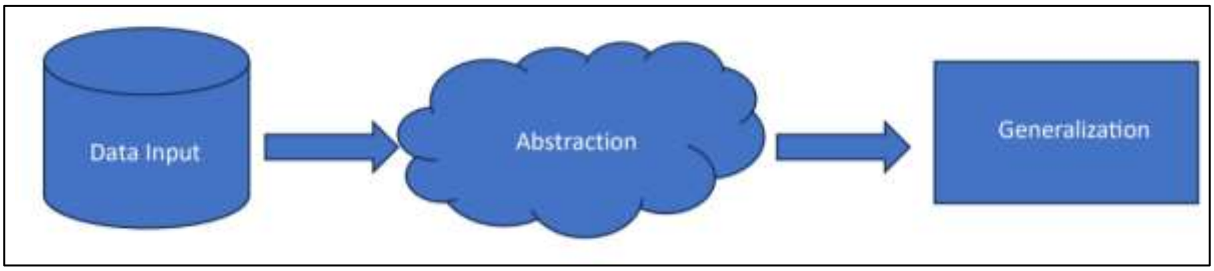

Figure 6 : basic machine learning process

The vast landscape of Machine Learning systems benefits from classification into groups based on key characteristics:

- **Supervision:** Does it need human-labeled data (supervised), explore itself (unsupervised), leverage a combination of labeled and unlabeled data (semi-supervised), or learn by doing (reinforcement)?)(Géron, 2022)
- **Learning Mode:** Does it adapt continuously (online) or in big chunks (batch)? (Géron, 2022)
- **Generalization Approach:** Does it compare new things to past examples (instance-based) or build general rules (model-based)? (Géron, 2022)

These characteristics are flexible and can be combined.

II.3 Types of Machine Learning Systems:

Machine learning systems can be categorized based on the level and nature of supervision they receive during the training process.

II.3.1 Supervised learning:

Supervised learning constitutes a branch of machine learning in which algorithms discern patterns between input data and desired outputs through labeled examples. The input dataset is typically partitioned into training and testing subsets. The training set is employed to train the model with a chosen algorithm, while the test set serves to assess the model's performance (Goodfellow et al., 2016).

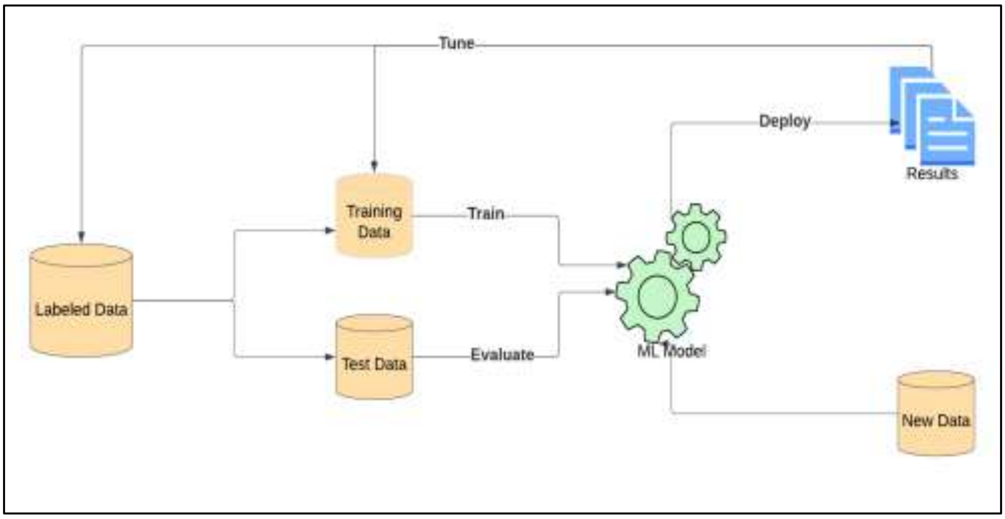

Figure 7 : Supervised learning Workflow.

Some of supervised learning algorithms are:

- **Decision Tree :**

A decision tree is a visual representation of choices and their outcomes, presented in the form of a tree-like structure. The nodes in the tree represent events or decisions, while the edges represent the conditions or rules that guide those decisions. The tree is made up of nodes and branches, with each node representing a group of attributes that need to be classified, and each branch representing a possible value that the node can take. By following the branches of the tree, one can navigate through the decision-making process and arrive at a final outcome or classification. (Mahesh, 2020)

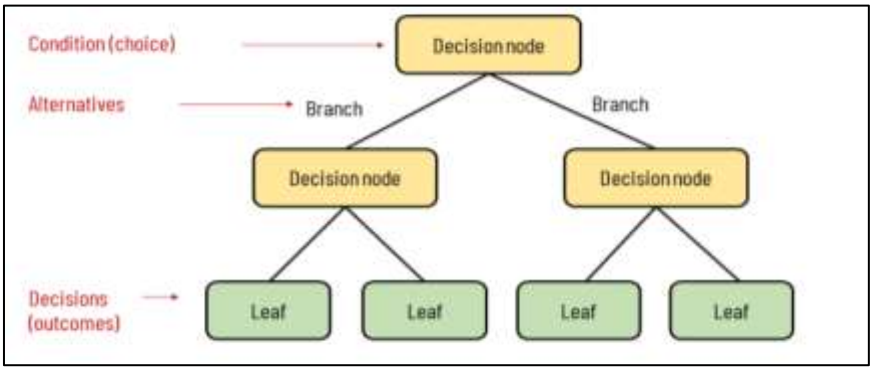

Figure 8 : Elements of a Decision tree

- **Naive Bayes :**

The Naive Bayes algorithm, based on Bayes Theorem, calculates the posterior probability (P(c|x)) given prior probabilities (P(c), P(x)), and the likelihood (P(x|c)). It assumes class conditional independence, meaning the effect of a predictor's value (x) on a class (c) is independent of other predictors' values. This method is particularly employed in the text classification industry, primarily for clustering and classification tasks based on conditional probability assessments (Mahesh, 2020).

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

$$P(c|X) = P(x_1|c) \times P(x_2|c) \times \ldots \times P(x_N|c)\, P(c)$$

- P(c|x) is the posterior probability of class (target) given predictor (attribute).
- P(c) is the prior probability of class.
- P(x|c) is the likelihood which is the probability of the predictor given class.
- P(x) is the prior probability of the predictor

- **SVM (Support Vector Machine) :**

SVM is a supervised learning algorithm used for data analysis in classification and regression tasks. It can handle both linear and non-linear classification through the kernel trick, which transforms input data into higher-dimensional feature spaces. SVMs establish margins between different classes, ensuring maximum distance between the margins and the classes to minimize classification errors. Widely applied across diverse domains including healthcare, natural language processing, signal processing, speech recognition, and image recognition, SVMs play a significant role in various applications (Mahesh, 2020).

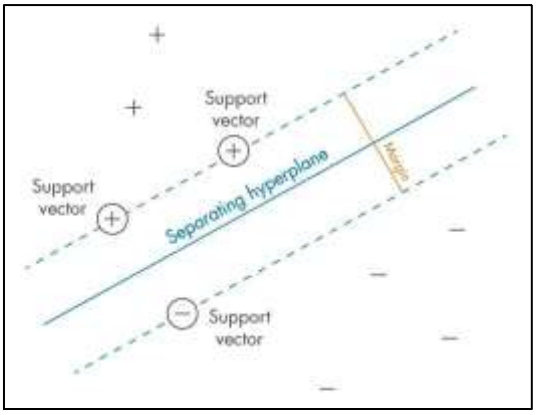

Figure 9 : SVM

II.3.2 Unsupervised Learning:

Unsupervised ML doesn't rely on predefined answers or a teacher's guidance. Instead, algorithms autonomously explore data to reveal hidden patterns and structures. They learn from data without specific labels, focusing on tasks like clustering and feature reduction. When faced with new data, they apply previously learned patterns to classify it (Goodfellow et al., 2016). Some of supervised learning algorithms:

- **Clustering:**

Clustering is a data mining technique that organizes unlabeled data into groups based on similarities or differences. It involves using clustering algorithms to group raw, unclassified data objects into structures or patterns. These algorithms fall into categories such as exclusive, overlapping, hierarchical, and probabilistic (Géron, 2022). Where exclusive clustering assigns each data point to just one cluster, overlapping clustering permits data points to belong to multiple clusters with varying degrees of membership, hierarchical clustering uses two main approaches: "bottom-up" (agglomerative) merging individual points and "top-down" (divisive) splitting a single cluster., and probabilistic clustering clusters data points based on their likelihood of belonging to a particular distribution.

- **Dimensionality Reduction:**

In the context of machine learning, while an increase in data volume generally enhances the accuracy of results, it can also negatively impact algorithm performance by causing issues such as overfitting and complicating data visualization. Dimensionality reduction is a technique employed when datasets contain an excessive number of features or dimensions. This process reduces the number of data inputs to a more manageable size, all while striving to maintain the dataset's integrity (Géron, 2022). It is particularly valuable during the data preprocessing stage. Several methods of dimensionality reduction include:

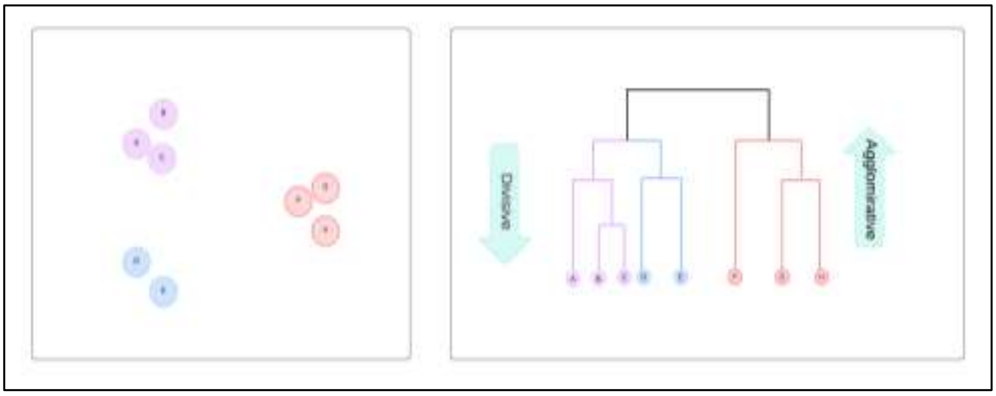

Figure 10 : Clustering

- **Principal Component Analysis (PCA):**

Principal Component Analysis is a widely-used dimensionality reduction algorithm that minimizes redundancies and compresses datasets through feature extraction. This method applies a linear transformation to the data, generating a set of "principal components." The first principal component identifies the direction that maximizes the dataset's variance. The second principal component, uncorrelated to the first, also finds maximum variance but in an orthogonal direction. This iterative process continues for subsequent components, each orthogonal to the previous ones, capturing the most variance possible in descending order.

- **Autoencoders:**

Autoencoders use neural networks for dimensionality reduction by compressing the data and subsequently reconstructing a new representation of the original input. In this framework, the hidden layer acts as a bottleneck, compressing the input data before it is reconstructed in the output layer. The process of transforming data from the input layer to the hidden layer is known as "encoding," and the transformation from the hidden layer to the output layer is called "decoding." This technique efficiently reduces dimensionality while preserving essential data features. (Géron, 2022).

These techniques are integral to enhancing the efficiency and performance of machine learning models by addressing the challenges posed by high-dimensional data.

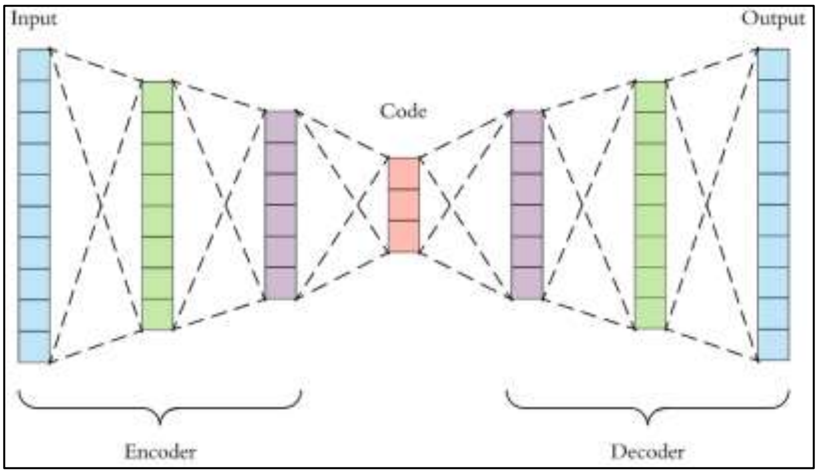

Figure 11 : Autoencoder (Dertat, 2017)

II.3.3 Semi Supervised Learning (SSL):

SSL is a machine learning method that trains a model using a small set of labeled data combined with a larger set of unlabeled data. This bridges the gap between supervised learning (needing a lot of labeled data) and unsupervised learning (not using labeled data at all) (Reddy et al., 2018). SSl is divided into two types:

- **Semi-Supervised Classification (SSC):** SSC classifies data using a mixture of a limited amount of labeled data and a substantial volume of unlabeled data, minimizing the requirement for costly and time-intensive labeled data generation (Reddy et al., 2018).
- **Semi-supervised clustering:** Semi-supervised clustering leverages both labeled data and unlabeled data, along with constraints (like "must-link" and "cannot-link" relationships), to guide the clustering process and group data points into meaningful clusters. Semi-supervised Single Link, a particular approach, tackles the issue of arbitrarily shaped clusters by leveraging a predefined distance matrix with minimal constraints. This method has been shown effective on both synthetic and real-world datasets by using the Self-training approach that consists of using a classification algorithm with few labeled training data, then it classifies an unlabeled data and these predicted patterns then added to the training set. The process is repeated until the test set is empty (Reddy et al., 2018).

II.3.4 Reinforcement Learning (RL):

RL is an ML technique where an agent interacts with its environment (e) through trial and error. The agent observes the environment, takes actions (A), to move from a state to another (S) and receives rewards or penalties (R) in return. The goal of the agent is to learn a policy, a strategy that dictates the best action to take in a given situation, to maximize its long-term reward (Akanksha et al., 2021)

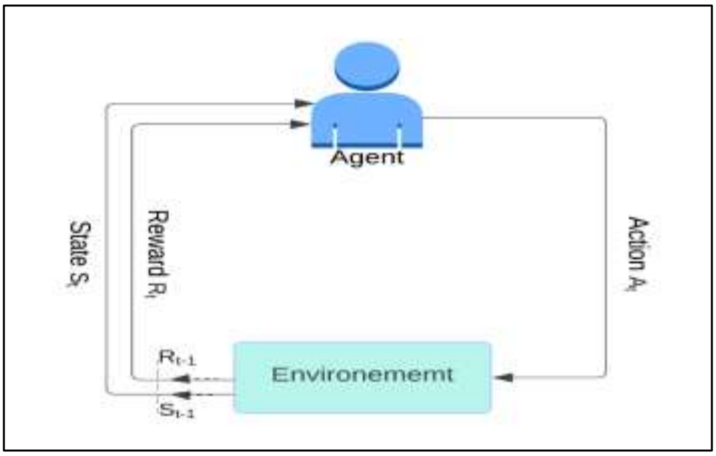

Figure 12 : Reinforcement learning

II.4 Learning Mode:

Another criterion for categorizing Machine Learning systems is their capability to learn incrementally from a continuous stream of incoming data (Géron, 2022).

II.4.1 Batch learning bottleneck:

Batch learning trains a machine learning model by feeding it the entire dataset at once. This offline process is computationally expensive and time-consuming, making it less ideal for real-time updates, large datasets, or resource-limited environments. Updating the model with new data requires retraining it from scratch using the complete dataset, including both old and new information. However, for systems requiring rapid adaptation to changing data or operating under resource constraints, batch learning may not be suitable. Incremental learning algorithms offer a more efficient alternative in such cases (Géron, 2022).

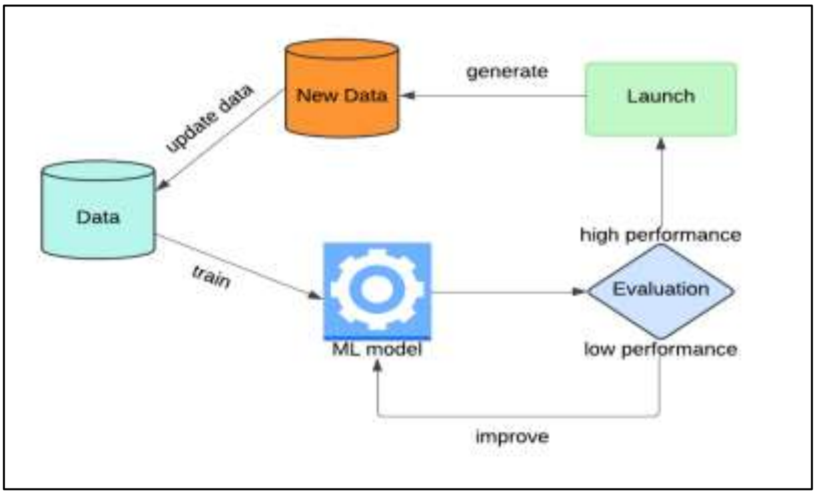

Figure 13: Batch leaning

II.4.2 Online Learning:

Online learning is a method where the system is trained incrementally by feeding it data instances sequentially, either individually or in small groups known as mini-batches. Each learning step is fast and economical, allowing the system to learn about new data as it arrives. This approach is particularly suitable for systems that receive data continuously and need to adapt quickly or autonomously, such as those dealing with stock prices. Additionally, online learning is advantageous when computing resources are limited because once the system learns from new data instances, it can discard them, saving space. Online learning algorithms can also handle huge datasets that cannot fit into a single machine's memory through a process called out-of-core learning. In out-of-core learning, the algorithm loads part of the data, performs a training step, and repeats this process until all data has been processed. A crucial parameter for this method is the learning rate, determining how swiftly the system adjusts to changing data. A high learning rate facilitates rapid adaptation to new data, yet risks overlooking older data. Conversely, a low learning rate fosters slower adaptation, but provides greater stability by reducing sensitivity to noise or outliers in the new data. However, a challenge with online learning is that if bad data is fed to the system, its performance may gradually decline (Géron, 2022).

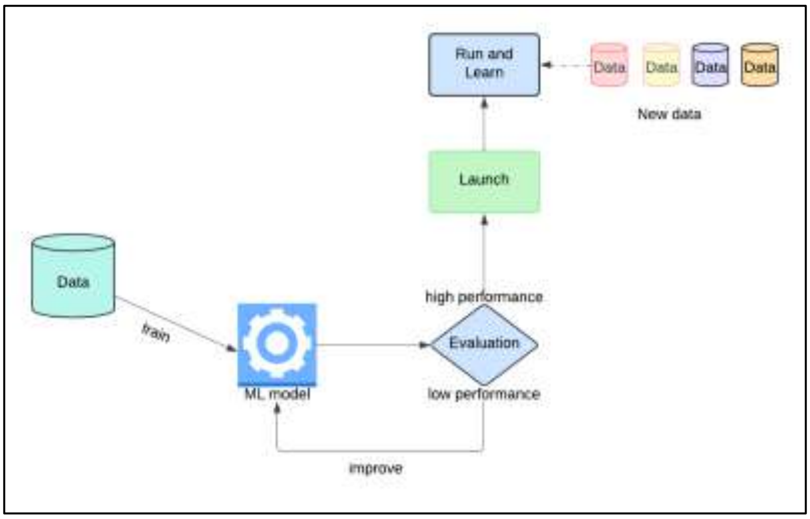

Figure 14 : Online learning process

II.5 Generalization approach:

ML algorithms can also be categorized based on their approach to generalization, which refers to their ability to make accurate predictions on unseen data. While training performance is important, the ultimate goal is for the model to perform well on entirely new examples. There are two main approaches to achieving generalization: instance-based learning and model-based learning (Géron, 2022).

II.5.1 Instance-Based learning:

Instance-based learning is a ML approach where the system memorizes examples and then generalizes to new cases by using a similarity measure to compare them to the learned examples or a subset of them. This method involves storing training instances and making predictions based on the similarity of new instances to these stored examples, without explicitly creating a generalized model (Géron, 2022).

II.5.2 Model-based learning:

Model-based learning involves creating a model that generalizes from the training data to make predictions on new, unseen data. The model captures patterns, relationships, and features within the data to provide a way to understand and predict outcomes (Géron, 2022).

Instance Based Learning	Model Based Learning
Does not create an explicit generalized model.	Creates a generalized model from the training data
Stores the training examples and compares new instances to them directly.	Make predictions on new, unseen instances.
Limited ability to generalize well to new, unseen data.	The model captures patterns and relationships in the data, allowing for better generalization.

Table 2 : Comparison between Model-Based vs Instance-Based

III. Deep Learning:

III.1 Definition:

“Deep learning is a subset of machine learning that uses multi-layered neural networks, called deep neural networks, to simulate the complex decision-making power of the human brain. Some form of deep learning powers most of the AI in our lives today.” (*What Is Deep Learning?*, 2021).

III.2 DL Architectures:

Here are some key deep learning architecture:

III.2.1 Artificial Neural Network (ANN):

ANNs are inspired by the brain and mimic its learning process. These networks consist of interconnected processing units called artificial neurons. Each connection between neurons has a weight that determines its influence. By adjusting these weights during training, ANNs learn the relationships between inputs and outputs. This allows them to tackle complex, non-linear problems. ANNs are popular due to their ease of use thanks to widely available libraries and their ability to handle new data beyond the training set, making them highly generalizable (Scalabrini Sampaio et al., 2019).

In Multilayer Perceptron (MLP), which is the base of artificial neural networks, the neurons are organized into three types of layers (Arruda et al., 2022):

- **Input Layer:** This layer receives the initial data or features that the neural network will process. Each neuron in the input layer represents one feature of the input data.
- **Hidden Layers:** These layers perform computations on the input data. Each neuron in a hidden layer receives input from neurons in the previous layer, applies a mathematical function (called an activation function), and passes the result to neurons in the next layer.

- **Output Layer:** The final layer of the neural network generates the model's output. The number of neurons in this layer depends on the type of problem the neural network is addressing. For instance, in a binary classification problem (where the output is either "yes" or "no"), there would be one neuron in the output layer.

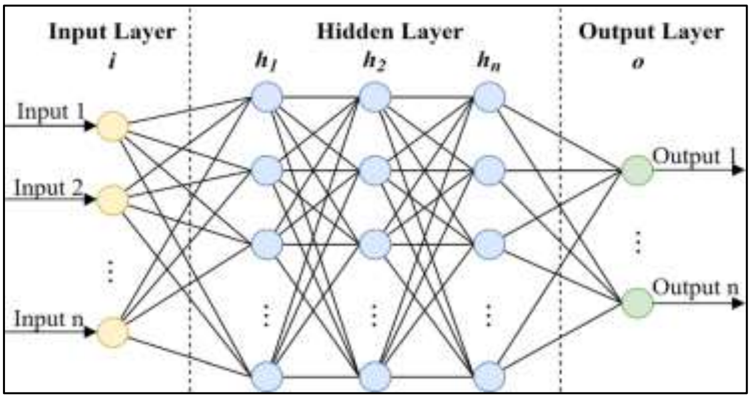

Figure 15 : General neural network architecture (Balakrishnan et al., 2022)

III.2.2 Convolutional neural networks (CNNs):

A Convolutional Neural Network (CNN) is a specific class of feed-forward ANN that integrates convolution operations into at least one of its layers. Inspired from the structure of biological neural networks, CNNs merge ANN principles with discrete convolution techniques primarily for image processing tasks, enabling automatic feature extraction. Consequently, CNNs are tailored for recognizing and analyzing two-dimensional data, such as images and videos, with the advantage of directly accepting images as input, thus eliminating the need for intricate feature extraction and data transformation procedures typically found in conventional image recognition algorithms (Gu et al., 2019).

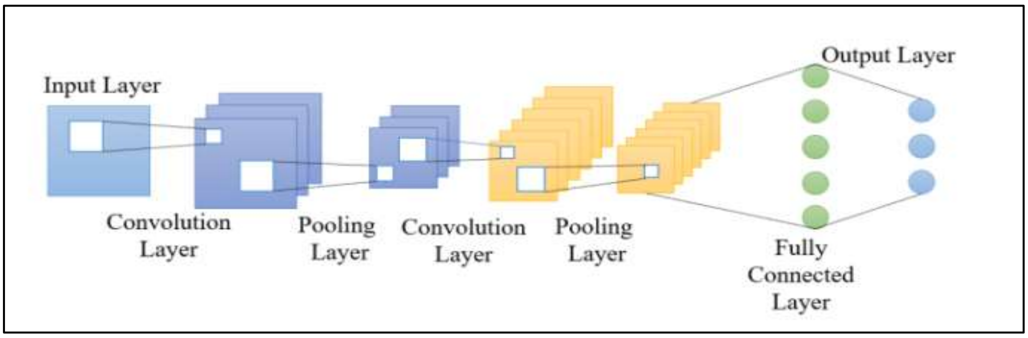

Figure 16 : Basic CNN architecture (Gu et al., 2019)

III.2.3 Recurrent neural networks (RNNs):

RNNs are a type of neural network particularly suited for processing sequential inputs like speech and language. Unlike traditional feedforward networks, RNNs process input sequences one element at a time while maintaining a "state vector" in their hidden units, which encodes information about the history of past elements in the sequence. This allows RNNs to capture temporal dependencies within the data (Feng et al., 2017).

In essence, RNNs can be conceptualized as deep multilayer networks where the outputs of hidden units at different time steps act like outputs of distinct neurons. This enables the application of backpropagation for training.

However, training RNNs has presented challenges due to the tendency of backpropagated gradients to either exponentially increase or decrease at each time step, leading to the problem of exploding or vanishing gradients over extended sequences (Feng et al., 2017; LeCun et al., 2015).

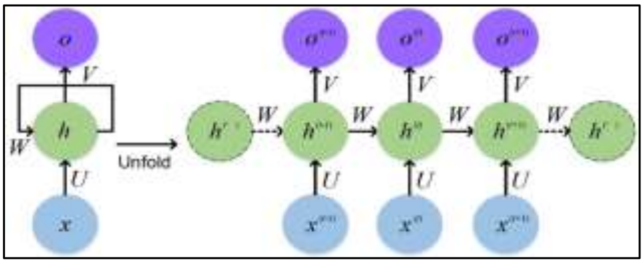

Figure 17 : The standard RNN and unfolded RNN (Feng et al., 2017)

III.2.4 Long short-term memory (LSTM):

Long Short-Term Memory (LSTM) is a type of recurrent neural network architecture that is well-suited for modeling sequential or time series data; they were introduced to overcome the vanishing gradient problem faced by traditional RNNs when training on long sequences (Elmaz et al., 2021; Goodfellow et al., 2016).

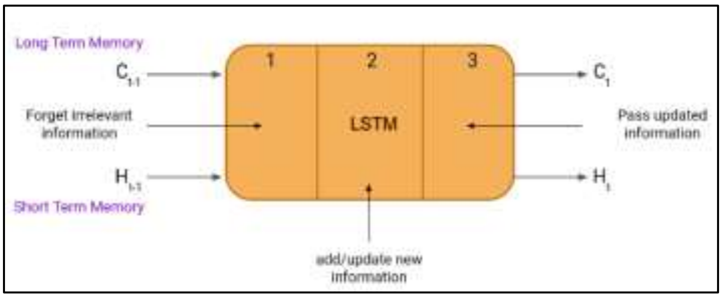

Figure 18 : Logic behind LSTM (analyticsvidhya)

IV. Data quality challenges in PdM:

In smart manufacturing, the adoption of machine learning for predictive maintenance has shown great promise in enhancing operational efficiency and product quality. However, the effectiveness of these predictive models heavily depends on the quality of the data used for training. As highlighted in the recent work by (Bharot et al., 2024), several data quality challenges can significantly impair the performance of predictive maintenance models:

IV.1 Missing Values:

Industrial datasets often contain missing values due to sensor malfunctions, data transmission errors, or gaps in data collection. These missing values can lead to incomplete information, making it difficult for machine learning algorithms to learn meaningful patterns effectively (Bharot et al., 2024).

IV.2 Outliers:

Extreme or anomalous data points, possibly resulting from sensor glitches or rare events, can disproportionately influence model training. If not properly handled, outliers can skew the analysis and lead to incorrect predictions, especially in sensitive tasks like faulty product detection (Bharot et al., 2024).

IV.3 Inconsistencies and Measurement Errors:

Inconsistencies in data formats or units, along with errors in data collection or sensor measurements, introduce noise into the dataset. This noise can obscure the underlying patterns that machine learning models aim to capture, leading to decreased accuracy and reliability (Bharot et al., 2024).

IV.4 Class Imbalance:

In predictive maintenance, the instances of faulty products (minority class) are typically far fewer than non-faulty products (majority class). This imbalance can result in biased models that excel at predicting the majority class but perform poorly on the critical minority class, failing to detect faulty products effectively (Bharot et al., 2024).

IV.5 High Dimensionality:

Manufacturing processes often generate high-dimensional data, with hundreds or even thousands of features. For instance, the SECOM dataset used by Bharot et al. contains 591 features. High dimensionality can lead to increased computational complexity and model performance deteriorates as the number of features grows (Bharot et al., 2024).

IV.6 Noisy Features:

Not all features in a dataset contribute meaningfully to the prediction task. These noisy or irrelevant features can obscure important patterns, leading to overfitting or increased training time without improving model performance (Bharot et al., 2024).

Addressing issues like missing values and class imbalance through data cleaning techniques significantly improves model performance, leading to better efficiency and reduced waste.

V. IoT and Sensors:

V.1 Definition of IoT:

IoT is a network of physical devices embedded with sensors, software, and other technologies that connect and exchange data with other devices and systems over the internet. This allows everyday objects to collect and share data, creating a more connected and automated world. So IoT is a bridge between the digital domain and the physical domain (Nguyen et al., 2019).

V.2 Industry Internet of things (IIoT):

The Industrial Internet of Things (IIoT) harnesses IoT technology to enhance industrial manufacturing processes. A defining feature of the IIoT is the integration of sensors into every component involved in manufacturing. These sensors act as the eyes of the system, collecting data throughout the manufacturing process and product lifecycle. Together with wireless sensor networks, they generate vast amounts of data, contributing to the emergence of Big Data in smart manufacturing systems (Nguyen et al., 2019).

The scale, complexity, and diversity of Big Data necessitate advanced computing technologies. AI now plays a crucial role in efficiently processing this data. Unlike traditional programming where computers execute specific tasks, AI imbues machines with intelligence, enabling them to interpret external data accurately, learn from it, and adapt flexibly to achieve various objectives (Nguyen et al., 2019).

The integration of AI has the potential to revolutionize industrial manufacturing processes across numerous applications. These include predictive quality analytics, automation, and insightful identification of engineering systems. By leveraging AI capabilities, manufacturers can enhance efficiency, optimize operations, and drive innovation in the industrial sector (Nguyen et al., 2019).

V.3 Big Data:

Within a smart factory environment, an extensive array of sensors is deployed for data collection. These sensors function by converting physical conditions into electrical signals, which are subsequently transmitted to a programmable logic controller for further processing. Each component within smart manufacturing possesses the ability to communicate and exchange data using cutting-edge network technologies. The rapid advancement of sophisticated IIoT technologies simplifies the process of data acquisition and storage, facilitating the emergence of the industrial Big Data era. Big data refers to massive, complex datasets with five key characteristics: volume (huge amounts of data), velocity (rapidly generated), variety (structured, unstructured, and semi-structured formats), veracity (accuracy, truthfulness, and trustworthiness of data.) and volatility (the rate of change and lifespan of data). In manufacturing, big data analytics unlocks the hidden value within this data. This process involves collecting, storing (often in the cloud), cleaning, analyzing, and visualizing the data (Nguyen et al., 2019).

By harnessing big data analytics, manufacturers can:

- **Optimize processes:** Identify and address inefficiencies for smoother production.
- **Predict and prevent equipment failures:** Reduce downtime and maintenance costs.
- **Improve product quality:** Analyze data to identify and fix quality issues early.
- **Gain insights from customer data:** Develop targeted marketing strategies and create new products.

Figure 19 : A framework for the smart factory

V.4 IoT for predictive maintenance:

IIoT, Big Data, and AI converge to enable predictive maintenance, a cornerstone of intelligent manufacturing. This approach, also known as just-in-time maintenance, revolutionizes traditional maintenance practices. In the past, maintenance occurred at fixed intervals, often resulting in lost productivity and unnecessary downtime. However, predictive maintenance harnesses real-time data from sensors embedded across equipment to continuously assess their health. Through AI-powered analytics, this data is processed to forecast potential failures or overloads, optimizing maintenance schedules. Advanced techniques such as Online Learning, Transfer Learning, and Domain Adaption further enhance predictive maintenance capabilities, aligning with the principles of Industry 4.0 (Nguyen et al., 2019).

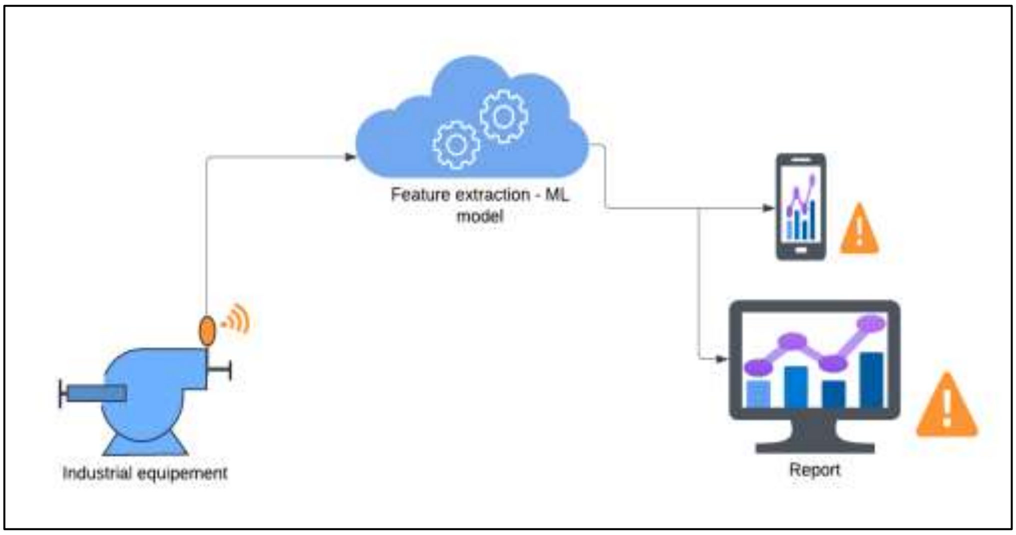

Figure 20 : IOT in Industry

VI. Conclusion:

In conclusion, this chapter explored the integration of machine learning algorithms, sensor technology, and the IIoT, a technological combination transforming predictive maintenance. By analyzing vast sensor data, ML, particularly deep learning, enables precise equipment health predictions. This collaboration optimizes performance, minimizes downtime, and maximizes operational efficiency, underscoring its critical role in advancing maintenance strategies. As these technologies continue to evolve, their impact on predictive maintenance will deepen, heralding a new era of innovation and excellence in the field.

Chapter 03: State of the art

I. Introduction:

This chapter delves into the domain of data-driven predictive maintenance, a critical strategy for ensuring system reliability and optimizing resource allocation. Here, we conduct a comprehensive literature review to explore the existing techniques and approaches in this field, with a specific focus on ML models. This review aims to:

- ➢ Summarize the various ML models employed for system health assessment in PdM.
- ➢ Compare and contrast these approaches, highlighting their strengths, limitations, and real-world applications.
- ➢ Identify key trends, challenges, and potential gaps in current research on ML for PdM. This will pave the way for further exploration and development in this domain.

II. Single model:

In the literature, single models for PdM have been developed and have shown good results in various cases. In the following sections, we will cite and discuss some notable examples.

II.1 Supervised model:

The study builds upon the work of (Falamarzi et al., 2019), who explored the use of ML to predict tram track widening (gauge deviation). The researchers focused on two common techniques: artificial neural networks (ANNs) and support vector regression (SVR). They used data from Melbourne's tram system, considering factors such as existing track width, track condition, and traffic patterns. ANN and SVR models were developed to predict gauge deviation for both straight and curved track segments, with the models trained on 75% of the data and tested on the remaining 25%.
The models performed well, particularly for straight tracks, achieving a coefficient of determination ($R^2 > 0.9$), and indicating high prediction accuracy. Curved tracks proved more challenging, with an ($R^2 > 0.75$), but still showed good potential. These results suggest that both ANN and SVR models can be effective in predicting tram track degradation.

The study indicates that ANNs might have a slight edge in predicting gauge deviation for straight tracks, while SVR may be more suitable for curved tracks. Cross-validation ensures model accuracy, with ANN excelling on straight tracks and SVR on curves, suggesting a combined model could handle complex datasets better, as (Saxena & Saad, 2007) used a combination of genetic algorithms (GA) and ANNs for condition monitoring of rotating mechanical systems like bearings (Saxena & Saad, 2007).

(Hesser & Markert, 2019) contributed to the advancement of intelligent tool wear monitoring by investigating the effectiveness of artificial neural networks (ANNs) in a retrofitted scenario. The study targeted CNC milling machines exceeding 15 years of age, aiming to bridge the gap and integrate them into Industry 4.0 principles. A Bosch XDK sensor was installed to capture vibration data during cutting processes using both brand new and worn-out tools for clear data contrast. Extracted features, including root-mean-square value, standard deviation, and area under the envelope signal, were derived from the raw vibration data to represent the tool's health. A supervised ANN model is trained using these extracted features to classify tool condition (brand-new vs. worn-out). The ANN demonstrates (high accuracy 0.9444 recall 0.9687 and precision 0.9393) in predicting tool states, outperforming SVM and KNN classifiers. This superior performance highlights the ANN's ability to effectively distinguish between healthy and worn tools based on subtle variations in vibration signatures.

Another study tackles motor failure prediction using ANNs. It leverages the data processing approach from (Scalabrini Sampaio et al., 2019), where simulated motor vibrations (created with a fan and magnets) are measured with an accelerometer. Extracted features like frequency and amplitude are obtained from the vibration signal. Additionally, an estimated time-to-failure is calculated based on a growth rate assumption. This processed data, along with the raw vibration measurements, forms the training dataset. The optimal ANN configuration achieved a learning rate of 0.85, trained for 50,000 epochs with a single hidden layer containing 25 neurons. Backpropagation was used for training, with logistic and linear activation functions applied in the hidden and output layers respectively. While the ANN performed well in cross-validation (RMSE: 0.0038), Random Forests achieved a slightly lower error (0.0026). However, the ANN excelled in real-world scenario tests (generalization), achieving significantly lower RMSEs (0.0313 and 0.1184) compared to other techniques like Random Forests, Regression Trees, and SVM. This showcases the ANN's strength in predicting remaining useful life, especially for longer timeframes.

The extraction of features such as root-mean-square value, standard deviation, area under the envelope signal, frequency, and amplitude from raw vibration data significantly enhances monitoring capabilities by providing a detailed representation of the tool's health, which allows the ANN to make more precise predictions. Additionally, the ANN model outperforms other techniques like Regression Tree and Random Forest, offering more effective predictions across short, medium, and long-term timeframes (Scalabrini Sampaio et al., 2019). However, the approach in (Hesser & Markert, 2019) is limited in its diagnostic scope, as it classifies the current state of tool wear but lacks prognostic capabilities to predict the remaining useful life or future failure states. Moreover, the methodology is not complete by considering only vibration data, whereas real-world equipment involves other critical variables such as temperature and pressure, necessitating methodological adjustments for accurate predictions.

Exploring the powerful feature extraction and learning capabilities of CNNs, (Ding et al., 2021) introduced a regression-based approach for predicting the RUL of bearings. The process begins by addressing gross monitoring errors in the original data using the 3-sigma criterion, which removes significant anomalies based on the distribution characteristics and acquisition methods of the monitoring data. After denoising, frequency features are extracted from the time-domain data using FFT, with RMS used as a tracking metric to depict the trend in bearing vibrations. To mitigate the influence of data fluctuations, these features are normalized. The dataset is then divided into training and testing sets through stratified sampling based on time and frequency features. A machine learning model is iteratively trained and validated to minimize error. To validate this approach, a case study was conducted using the NASA IMS bearing dataset.

The study involved three experiments. In the first experiment, the dataset of bearing 1 was used with various algorithms: RNN, LSTM, W-CNN (WINDOW-CNN), N-DCNN (Normal DCNN), and DCNN (Deep CNN) were tested without denoising, with Sliding median denoising, hard threshold denoising, Soft threshold denoising, Singular value decomposition (SVD) denoising, and 3-sigma criterion denoising. The best results were obtained by using 3-sigma criterion denoising combined with DCNN. The second experiment evaluated the generalization ability of the proposed model by using the datasets of bearings 1, 2, and 3 for testing all the previous models with 3-sigma criterion denoising. Again, DCNN showed the best generalization. The final experiment examined the validity of stratified sampling. To verify the influence of stratified sampling on the prediction results, the dataset was divided according to the traditional time series data partitioning method, based on the ratio 8:2, while keeping the denoising method and the prediction model unchanged. The dataset of four bearings was applied, and the results showed that stratified sampling was superior.

This approach is rooted in extensive empirical research and experimentation, demonstrating its robustness and effectiveness through comprehensive testing of various combinations. The development process involved methodically evaluating multiple algorithms and denoising methods that not only highlights the method's robustness but also ensures its adaptability and effectiveness in predicting the RUL of bearings. However, the testing of the generalization method was based only on the data of the same experience.

II.2 Unsupervised:

(Von Birgelen et al., 2018), proposed an unsupervised data-driven approach using Self-Organizing Maps for anomaly detection, localization, and predictive maintenance in Cyber-Physical Production Systems (CPPS). The approach leverages a SOM trained on data representing the system's normal condition to establish a baseline for normal behavior. Anomalies are detected by calculating the quantization error (distance between a data sample and its best matching unit in the SOM and comparing it to a predefined threshold. Anomaly localization is achieved by identifying the signals exhibiting the largest deviation from the expected values predicted by the SOM. The system's degradation over time is monitored by tracking the increase in quantization errors and pinpointing the probable cause of degradation, facilitating informed maintenance decisions. The effectiveness of the proposed approach is demonstrated through testing on a benchmark bearing dataset, where it successfully detected degradation and identified the failing bearing before complete failure. Furthermore, the approach was applied to two real-world industrial use cases: Predicting degradation of a critical component at Reifenhäuser Reicofil based on multiple sensor signals and monitoring degradation of cutting blades in shrink-wrapping machines at Ocme by comparing data from new and worn blades.

In both cases, the data-driven SOM approach successfully detected and localized degradation within the real-world industrial systems. Moreover, the study highlights the transferability of the learned SOM model, demonstrating its applicability across various systems and datasets. While the approach excels at detecting anomalies, estimating the RUL of components requires further development.

II.3 Traditional machine learning vs. Deep learning models:

(Namuduri et al., 2020)presented a case study by the sensors group at Florida International University in collaboration with the sensors and systems group at University of Dayton Research Institute used a dataset that consists of data from 21 sensors on a turbofan aircraft engine which is a part of a NASA's prognostics data repository. These sensors measure various parameters, including temperature, pressure, fan speed, fuel-air ratio, and coolant bleed. The authors focused on a binary classification task, predicting failures within a cycle based on sensor data. Five algorithms were tested: LR, SVM, Ensemble Model, ANN, and LSTM. LSTM achieved the best results with an accuracy of 99.3% and AUC of 0.998 and a superior precision 87.3% compared to other models, demonstrating its effectiveness in learning from sequential sensor data.

(Silvestrin et al., 2019), used data comprises a collection of sensors installed in a hydraulic test rig. Two deep learning architectures, TCN and LSTM, were employed, both using the Adam optimizer with an initial learning rate of 0.01 for 200 epochs. The objective function for both architectures is cross-entropy, which quantifies the disparity between predicted and actual class probabilities. In the LSTM model, hidden units are employed for memory, and it directly outputs class probabilities. Conversely, the TCN model extracts features using dilated convolutions and subsequently predicts class probabilities. The traditional machine learning algorithms including KNN, DT, and RF outperformed complex Deep Learning models especially the tree based models due to its efficiency in learning from smaller datasets.

Choosing between traditional ML and DL for PdM requires consideration of data availability and task complexity. Deep learning excels with large datasets, offering superior predictive performance due to its effectiveness in learning from sequential sensor data (Namuduri et al., 2020). However, its limited and small data can be a challenge. Traditional ML methods, on the other hand, are more data-efficient and provide interpretable insights. While they might struggle with complex tasks involving sequential data, they can outperform deep learning with limited examples, especially the tree based model (Silvestrin et al., 2019). Regardless of the chosen approach, sensor design plays a critical role in ensuring high-quality data. Future research should explore validating these findings across a wider range of datasets and tasks. A promising direction could involve developing hybrid approaches that leverage the strengths of both ML and DL for even more effective PdM solutions.

III. Multi - Models:

Multimodel approaches are more frequently employed for predictive maintenance (PdM) and have consistently delivered better results across various applications. In the following sections, we will highlight and discuss some notable examples.

III.1 Statistical model stacked with Supervised ML algorithms:

The ARIMA model functions as a powerful tool for forecasting machine parameters by scrutinizing trends within sensor-collected data, thus facilitating the anticipation of future states of the machinery. Its utility extends to predictive maintenance, wherein it assists in projecting forthcoming data points indicative of potential machine conditions. This application of ARIMA enables the implementation of proactive maintenance strategies rooted in predictive insights derived from the analysis of historical sensor data.

(Kanawaday & Sane, 2017) gathered data as a time series from slitting machines via sensors and transmitted to the cloud in CSV format. When faced with new production cycles, the ARIMA model generates forecasts for parameter values throughout the cycle's duration. These forecasted values are then leveraged by the supervised model, which may include algorithms such as Naive Bayes, SVM, Classification and Regression Trees (CART), and Deep Neural Networks (DNN). Notably, the DNN model achieved the highest prediction accuracy of 98.69%.

(Francis & Mohan, 2019) included the implementation of an automated data capture system in railway transportation that offers insights into fault trends and enables failure prediction. This system is divided into training and prediction phases. During training, past event data undergoes ARIMA-based feature extraction, followed by PCA for feature reduction. These features are then trained with ML algorithms (RF). For new events, PCA features are analyzed by a SVR model to decide on maintenance necessity.

Analyzing trends with ARIMA models followed by separate failure prediction models is a good starting point for PdM, but it lacks crucial RUL estimation. While it effectively predicts failures, knowing the estimated time a machine has left before failure allows for proactive maintenance scheduling, optimizing resources. Incorporating an additional RUL prediction model could significantly enhance this framework's effectiveness.

III.2 Two stages data-driven framework:

(Xiang et al., 2018) presented a data-driven framework for diagnosing and predicting issues in complex vending machines. By leveraging various machine logs, the framework automates fault diagnostics and predicts component failures through feature extraction and supervised learning with precise data labeling. This approach, built upon the work of (Xiang et al., 2018) achieves high accuracy (>80% in precision, recall, and F-measure), showcasing its potential to revolutionize maintenance efficiency and significantly reduce costs. A key strength lies in the framework's data labeling methodology for supervised learning. By comparing serial numbers of core components on adjacent dates, the system accurately identifies failures (yA and yB) for Modules A and B. For prognostics, the framework addresses challenges in labeling the normal machine state. It effectively captures this state by labeling the middle point between failures as normal leading to more accurate predictions.

This two-stage framework with precise data labeling significantly outperforms traditional models, boosting prediction accuracy for various failure types (65.45% vs. 27.53%). This approach allows for prioritizing maintenance based on predicted machine health, but requires extensive labeled data and carries a risk of false positives/negatives due to its computational intensity.

III.3 Machine learning models (unsupervised and supervised):

An innovative approach was implemented by (Rousopoulou et al., 2019) to mitigate potential oven failures within the printed circuit board (PCB) production line at Boston Scientific, aiming for early detection before any significant disruptions occur. Five acoustic sensors were strategically placed throughout the production environment to continuously capture sound data emanating from the ovens. The system uses advanced outlier detection algorithms including Mean Absolute Deviation (MAD), Local Outlier Factor (LOF), and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) to analyze the collected acoustic data and identify potential oven failures and assess deviations from normal patterns within the data. By identifying outliers, the system can pinpoint anomalies that might signal impending oven failures. To address imbalanced data, where normal cases outnumber failures (86% of the dataset), and the SMOTE technique is employed. A minority class point is chosen randomly, and the k-neighbors of it are calculated. The synthetic data are placed between the picked point and the calculated neighbors. Leveraging machine learning capabilities, an SVM model is trained with the following parameters: cost = 0.5 and kernel = RBF to accurately predict failures, achieving a precision rate of 76%, an F1-Score of 86%, and a recall of 100%.

This approach tackles imbalanced data by generating synthetic data for failures (minority class). It then uses an SVM model to effectively classify normal and anomaly data. However, its effectiveness is limited by focusing on a specific data type (acoustic data from ovens) and requires expertise for parameter tuning in both outlier detection and the SVM model. Additionally, outlier detection algorithms might raise false positives, requiring further analysis and potentially combining them with other predictive techniques for a more robust solution.

III.4 Multiple Classifier approach using SVM and KNN:

The key concept of MC PdM revolves around the use of multiple classifiers, each focusing on a distinct failure horizon, to address the limitations inherent in single classifier approaches in predictive maintenance. By assigning classifiers to specific classification problems based on varying failure horizons, MC PdM optimizes cost considerations according to prevailing operating conditions.

In a specific application scenario for predictive maintenance in semiconductor manufacturing (Susto et al., 2014) proposed a method to address imbalanced data in semiconductor filament replacement (ion implantation), tackles the issue of limited failure data. It accomplishes this by considering data points preceding failures as positive examples. The classification algorithms used in this case include k-NN and SVM algorithms. SVM implemented with RBF kernels to create non-linear classification boundaries is useding, along with linear SVM that uses its capability to measure distance from the decision boundary to assess a sample's proximity to potential failure. A shorter distance indicates a higher likelihood of impending failure enabling proactive maintenance interventions. The choice of the threshold affects the results. It has also been demonstrated in this case study that SVMs offer superior performance to k-NN classifiers when implementing MC PdM, and that in general, MC-PdM-knn consistently outperforms PdM approaches.

MC PdM boost prediction accuracy by targeting different failure stages and handling imbalanced data effectively. The study below shows that MC PdM using SVMs outperforms single-classifier approaches. However, managing multiple models increases complexity, SVM performance relies on specific threshold, and training costs can be high, requiring careful evaluation for practical use.

III.5 Physics based, statistical and unsupervised ML models:

(Amruthnath & Gupta, 2018) used vibration data collected every 240 minutes over 12 days, sampled at a frequency of 2048 Hz on both X and Y axes. Different features were extracted from this data, including peak acceleration, peak velocity, turning speed, RMS velocity, and damage accumulation. Initially, they hypothesized two data states: healthy and unhealthy. Principal Component Analysis (PCA) combined with the T^2 statistic allowed them to differentiate between these states 31 observations before the fault occurred, compared to only 11 observations using visual inspection of data plots alone. Further exploration involved unsupervised clustering algorithms using the elbow method and the nbClust package, three optimal clusters were determined. Applying hierarchical clustering, K-means, and C-means resulted in nearly identical cluster formations. Leveraging prior knowledge of the data, labels were assigned to each cluster: healthy, warning and faulty. These algorithms rely on distance metrics. For the final model, a Gaussian Finite Mixture Model (GFMM) was employed with the Expectation-Maximization (EM) algorithm. Unlike predefining cluster numbers, GFMM discovers optimal clusters and classifies data accordingly. While the model identified five components, closer examination revealed overlap between components 1 & 2 and 3 & 4. Reorganizing these components yielded a pattern very similar to the previous cluster analysis.

This approach boasts earlier fault detection, robust clustering, and flexible classification using PCA, T^2, and Gaussian Finite Mixture Models. It can identify faults much earlier than visual inspection and leverages various techniques to ensure consistent and reliable fault identification. Additionally, the model can automatically discover clusters without pre-defining their number, enhancing flexibility. However, interpreting initial clustering results and managing the multi-step process can be complex, significant computational resources are required, especially for real-time applications with high-frequency data and the effectiveness of this approach heavily relies on expert knowledge for accurate feature extraction and ongoing model calibration is necessary for long-term accuracy.

III.6 Knowledge based models (rule-based/Case-Based) and ML approach:

A novel framework for predictive building maintenance is proposed by (Cheng et al., 2020), integrating Building Information Modeling (BIM) and IoT technologies. The framework functions through two layers. The Information Layer facilitates seamless data exchange between BIM models, IoT sensor networks, and the Facilities Management (FM) system. The Application Layer houses four key modules specifically designed for PdM: Condition

Monitoring & Fault Alarming, Condition Assessment, Condition Prediction, and Maintenance Planning. These modules use sensor data, BIM information, and historical maintenance records to continuously monitor equipment health, predict future problems, and schedule proactive maintenance activities. From these combined datasets, relevant variables are selected and used to train ML algorithms: ANN and SVM. The data is then split into three sets: the train set (80%), the cross-validation set (10%), and the test set (10%). SVM appears to be a more suitable choice for PdM of MEP components in this study due to its superior accuracy (96.547% for SVM compared to 96.422% for ANN), lower error rates (SVM exhibited lower error rates across all metrics, with the mean absolute error, root mean squared error, and absolute average error for ANN being 3.578% larger than those for SVM, which were 3.427%), and faster processing time (ANN requires twice the time (0.21 s and 0.14 s) compared to SVM (0.09 s and 0.05 s)).

The combination of static building data (BIM), real-time sensor information (IoT), and historical maintenance records (FM), offers a comprehensive picture of a system's health. This integration of diverse data sources, along with ML algorithms, improves the accuracy and reliability of maintenance decisions. However, building these models can be hindered by a lack of readily available domain expertise. Furthermore, the complexity of integrating various data sources and managing multiple models increases the difficulty of implementation and maintenance. Lastly, the current model requires training for each new type of building component, limiting its scalability for widespread use.

IV. Comparative table of the different case studies:

Ref	Model	case study	Advantage	Model type
(Von Birgelen et al., 2018)	SOM	anomaly detection, localization for bearings degradation	showed good results for another case study (cutting blades degradation)	Single unsupervised ML
(Falamarzi et al., 2019)	ANN & SVR	predict tram track widening (gauge deviation)	ANN for straight tracks SVM for curve tracks.	Single supervised

(Namuduri et al., 2020)	**LSTM** & ANN, LR, SVM, EM	classification of Nasa's turbofan engine for PdM	effective for sequential sensor data	Single supervised DL
(Silvestrin et al., 2019)	**TCN** & **LSTM** KNN, DT, RF	multi classification on hydraulic test rig	DL good for large data ML in small dataset	Single supervised DL & ML
(Scalabrini Sampaio et al., 2019)	**ANN** & RF, RT, SVM	motor failure prediction + RUL estimation	ANN's strength in predicting remaining useful life, especially for longer timeframes.	Single supervised ML
(Hesser & Markert, 2019)	**ANN &** SVM	classify tool conditions (blades) to integrate old CNC milling machines to industry 4.0	effectively distinguish between healthy and worn tools based on subtle variations in vibration signatures	Single supervised ML
(Ding et al., 2021)	denoising +DCNN	predicting RUL of bearings	models showed better results with denoised data	Single supervised DL
(Francis & Mohan, 2019)	ARIMA +RF	predict failure in railway transportation	forecasts for parameter values with ARIMA	Multi Statistical + supervised ML
(Xiang et al., 2018)	ML models	diagnosing and predicting issues in complex vending machines	The use of 2 stage framework for diagnostic & prognostic task	Multi Supervised ML
(Rousopoulou et al., 2019)	anomaly detection (MAD+LOF+DBSCAN)+Smote(KNN)+ SVM	oven failure detection within the printed circuit board (PCB) production line	handles unlabeled and imbalanced data	Multi unsupervised + supervised ML

(Susto et al., 2014)	Multi classifier(SVM & KNN)	predictive maintenance in semiconductor manufacturing	imbalanced data handling, MC-SVM testing MC-knn generalization	Multi supervised ML
(Amruthnath & Gupta, 2018)	**GFMM & EM** + k-means, c-means	identify faults in vibration data	automatically discover clusters without pre-defining their number	Multi Statistical, k-based & unsupervised ML
(Cheng et al., 2020)	Data integration + **SVM** & ANN	predictive maintenance within Facilities Management system	provides a holistic system health assessment, enhanced by ML	Multi Knowledge based + unsupervised ML
(Saxena et al., 2007)	Genetic algorithms + ANN	condition monitoring for bearings	handle complex datasets	Multi stochastic + supervised ML

Table 3 : Comparative table of the case studies

V. Conclusion:

Single ML models, while useful for basic tasks in PdM, struggle to handle the complexities of real-world industrial data. To address this, researchers are turning to multi-model approaches that combine the strengths of different techniques. Physics-based models offer deep system understanding, statistical techniques extract key data features, and ML algorithms excel at pattern recognition and prediction. By working together, these models can capture the intricate patterns hidden within complex data. This leads to more accurate anomaly detection, fault classification, and RUL prediction leading to a more comprehensive robust and explicable PdM system.

General Conclusion

PdM has significantly evolved within Industry 4.0, leveraging ML models to enhance system health assessment and failure prediction. Single models such as ANNs and SVM have demonstrated effectiveness in specific tasks but often struggle with complex datasets and the intricate nature of real-world systems. For instance, ANNs used for tram track gauge deviation prediction show high accuracy but require the incorporation of additional factors such as temperature and load conditions to handle complex datasets effectively. Similarly, SVMs are proficient in distinguishing between normal and abnormal states but face scalability and computational efficiency challenges with large, heterogeneous datasets. However, single models can still be highly effective with the right input data. By focusing on feature engineering and selection to improve data quality and incorporating domain knowledge, such as industry-specific laws and principles, we can extract meaningful features and patterns, enabling single models to better handle complex datasets and leverage their strengths.

Multi-model and hybrid approaches have emerged as more comprehensive solutions, offering improved accuracy and reliability, particularly when combining physics-based models with statistical techniques and ML algorithms. These approaches are especially effective in capturing complex patterns and trends within data, improving the robustness and accuracy of PdM systems. For example, physics-based models provide a deep understanding of system behavior and underlying failure mechanisms. When combined with statistical models, they facilitate precise feature extraction, enhancing the quality of data fed into ML algorithms. Multi-model frameworks can detect anomalies more effectively by leveraging the strengths of different models. For instance, clustering algorithms can identify normal operating states, while supervised ML models can classify detected anomalies. Hybrid approaches using DL models like Convolutional Neural Networks and LSTM networks can predict RUL with high accuracy, benefiting from the rich feature sets derived from physics-based models.

Future researchers should explore further integration of different approaches and models to leverage their strengths and build a more robust framework. Hybrid approaches that combine DL with physics-based knowledge show particular promise in improving fault detection and RUL estimation.

The findings from this comparative study will set the stage for the practical component of our research, which will be outlined in a separate report. This practical phase will delve deeper into the implementation of a specific PdM approach within an industrial setting, leveraging the insights gained from the literature review and objective analysis. This integration of theoretical insights with practical application will provide a comprehensive understanding of PdM strategies and efficacy.

Bibliography:

Achouch, M., Dimitrova, M., Ziane, K., Sattarpanah Karganroudi, S., Dhouib, R., Ibrahim, H., & Adda, M. (2022). On predictive maintenance in industry 4.0: Overview, models, and challenges. *Applied Sciences*, *12*(16), 8081.

Adams, S., Malinowski, M., Heddy, G., Choo, B., & Beling, P. A. (2017). The WEAR methodology for prognostics and health management implementation in manufacturing. *Journal of Manufacturing Systems*, *45*, 82–96.

Akanksha, E., Sharma, N., & Gulati, K. (2021). Review on reinforcement learning, research evolution and scope of application. *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, 1416–1423.

Amruthnath, N., & Gupta, T. (2018). A research study on unsupervised machine learning algorithms for early fault detection in predictive maintenance. *2018 5th International Conference on Industrial Engineering and Applications (ICIEA)*, 355–361.

Arcos Jiménez, A., Gómez Muñoz, C. Q., & García Márquez, F. P. (2017). Machine learning for wind turbine blades maintenance management. *Energies*, *11*(1), 13.

Aremu, O. O., Palau, A. S., Parlikad, A. K., Hyland-Wood, D., & McAree, P. R. (2018). Structuring data for intelligent predictive maintenance in asset management. *IFAC-PapersOnLine*, *51*(11), 514–519.

Arruda, H. F. de, Benatti, A., Comin, C. H., & Costa, L. da F. (2022). Learning deep learning. *Revista Brasileira de Ensino de Física*, *44*, e20220101.

Ayo-Imoru, R. M., & Cilliers, A. C. (2018). Continuous machine learning for abnormality identification to aid condition-based maintenance in nuclear power plant. *Annals of Nuclear Energy*, *118*, 61–70.

Baban, M., Baban, C. F., & Moisi, B. (2018). A fuzzy logic-based approach for predictive maintenance of grinding wheels of automated grinding lines. *2018 23rd International Conference on Methods & Models in Automation & Robotics (MMAR)*, 483–486.

Balakrishnan, V., Shi, Z., Law, C. L., Lim, R., Teh, L. L., & Fan, Y. (2022). A deep learning approach in predicting products' sentiment ratings: A comparative analysis. *The Journal of Supercomputing*, *78*(5), 7206–7226.

Bengtsson, M. (2004). *Condition based maintenance systems: An investigation of technical constituents and organizational aspects*. Mälardalen University Eskilstuna, Sweden.

Berecibar, M., Devriendt, F., Dubarry, M., Villarreal, I., Omar, N., Verbeke, W., & Van Mierlo, J. (2016). Online state of health estimation on NMC cells based on predictive analytics. *Journal of Power Sources*, *320*, 239–250.

Bharot, N., Verma, P., Soderi, M., & Breslin, J. G. (2024). DQ-DeepLearn: Data Quality Driven Deep Learning Approach for Enhanced Predictive Maintenance in Smart Manufacturing. *Procedia Computer Science*, *232*, 574–583.

Cheng, J. C., Chen, W., Chen, K., & Wang, Q. (2020). Data-driven predictive maintenance planning framework for MEP components based on BIM and IoT using machine learning algorithms. *Automation in Construction*, *112*, 103087.

Climente-Alarcon, V., Nair, D., Sundaria, R., Antonino-Daviu, J. A., & Arkkio, A. (2017). Combined model for simulating the effect of transients on a damaged rotor cage. *IEEE Transactions on Industry Applications*, *53*(4), 3528–3537.

Culot, G., Nassimbeni, G., Orzes, G., & Sartor, M. (2020). Behind the definition of Industry 4.0: Analysis and open questions. *International Journal of Production Economics*, *226*, 107617.

Dalzochio, J., Kunst, R., Pignaton, E., Binotto, A., Sanyal, S., Favilla, J., & Barbosa, J. (2020). Machine learning and reasoning for predictive maintenance in Industry 4.0: Current status and challenges. *Computers in Industry*, *123*, 103298.

Ding, H., Yang, L., Cheng, Z., & Yang, Z. (2021). A remaining useful life prediction method for bearing based on deep neural networks. *Measurement*, *172*, 108878.

Downey, A., Lui, Y.-H., Hu, C., Laflamme, S., & Hu, S. (2019). Physics-based prognostics of lithium-ion battery using non-linear least squares with dynamic bounds. *Reliability Engineering & System Safety*, *182*, 1–12.

Elmaz, F., Eyckerman, R., Casteels, W., Latré, S., & Hellinckx, P. (2021). CNN-LSTM architecture for predictive indoor temperature modeling. *Building and Environment*, *206*, 108327.

Falamarzi, A., Moridpour, S., Nazem, M., & Cheraghi, S. (2019). Prediction of tram track gauge deviation using artificial neural network and support vector regression. *Australian Journal of Civil Engineering*, *17*(1), 63–71.

Feng, W., Guan, N., Li, Y., Zhang, X., & Luo, Z. (2017). Audio visual speech recognition with multimodal recurrent neural networks. *2017 International Joint Conference on Neural Networks (IJCNN)*, 681–688.

Francis, F., & Mohan, M. (2019). Arima model based real time trend analysis for predictive maintenance. *2019 3rd International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 735–739.

Géron, A. (2022). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. O’Reilly Media, Inc.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.

Gu, H., Wang, Y., Hong, S., & Gui, G. (2019). Blind channel identification aided generalized automatic modulation recognition based on deep learning. *IEEE Access*, *7*, 110722–110729.

Hameed, Z., Hong, Y. S., Cho, Y. M., Ahn, S. H., & Song, C. K. (2009). Condition monitoring and fault detection of wind turbines and related algorithms: A review. *Renewable and Sustainable Energy Reviews*, *13*(1), 1–39.

Hesser, D. F., & Markert, B. (2019). Tool wear monitoring of a retrofitted CNC milling machine using artificial neural networks. *Manufacturing Letters*, *19*, 1–4.

Hinchi, A. Z., & Tkiouat, M. (2018). Rolling element bearing remaining useful life estimation based on a convolutional long-short-term memory network. *Procedia Computer Science*, *127*, 123–132.

Hong, J., Miao, X., Han, L., & Ma, Y. (2009). Prognostics model for predicting aero-engine bearing grade-life. *Turbo Expo: Power for Land, Sea, and Air*, *48821*, 639–647.

Hu, C., Youn, B. D., Wang, P., & Yoon, J. T. (2012). Ensemble of data-driven prognostic algorithms for robust prediction of remaining useful life. *Reliability Engineering & System Safety*, *103*, 120–135.

Hussain, A., Lee, S.-J., Choi, M.-S., & Brikci, F. (2015). An expert system for acoustic diagnosis of power circuit breakers and on-load tap changers. *Expert Systems with Applications*, *42*(24), 9426–9433.

Industrial Revolutions. (n.d.). *NIST*. Retrieved June 3, 2024, from https://www.nist.gov/image/industrial-revolutions

Jimenez, J. J. M., Schwartz, S., Vingerhoeds, R., Grabot, B., & Salaün, M. (2020). Towards multi-model approaches to predictive maintenance: A systematic literature survey on diagnostics and prognostics. *Journal of Manufacturing Systems*, *56*, 539–557.

Kanawaday, A., & Sane, A. (2017). Machine learning for predictive maintenance of industrial machines using IoT sensor data. *2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, 87–90.

Kerkyras, S., Karakassidis, V., & Papaelias, M. (n.d.). *Condition monitoring of wind turbine gearboxes using acoustic emission*.

Kinghorst, J., Geramifard, O., Luo, M., Chan, H.-L., Yong, K., Folmer, J., Zou, M., & Vogel-Heuser, B. (2017). Hidden Markov model-based predictive maintenance in semiconductor manufacturing: A genetic algorithm approach. *2017 13th IEEE Conference on Automation Science and Engineering (CASE)*, 1260–1267.

Kumar, P., & Srivastava, R. K. (2012). An expert system for predictive maintenance of mining excavators and its various forms in open cast mining. *2012 1st International Conference on Recent Advances in Information Technology (RAIT)*, 658–661.

Lantz, B. (2019). *Machine learning with R: expert techniques for predictive modeling*. Packt publishing ltd.

Le Son, K., Fouladirad, M., Barros, A., Levrat, E., & Iung, B. (2013). Remaining useful life estimation based on stochastic deterioration models: A comparative study. *Reliability Engineering & System Safety*, *112*, 165–175.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444.

Li, Y., Liu, S., & Shu, L. (2019). Wind turbine fault diagnosis based on Gaussian process classifiers applied to operational data. *Renewable Energy*, *134*, 357–366.

Liao, S.-H. (2005). Expert system methodologies and applications—A decade review from 1995 to 2004. *Expert Systems with Applications*, *28*(1), 93–103.

Lu, F., Wu, J., Huang, J., & Qiu, X. (2019). Aircraft engine degradation prognostics based on logistic regression and novel OS-ELM algorithm. *Aerospace Science and Technology*, *84*, 661–671.

Ma, G., Jiang, L., Xu, G., & Zheng, J. (2015). A model of intelligent fault diagnosis of power equipment based on CBR. *Mathematical Problems in Engineering*, *2015*(1), 203083.

Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR).[Internet]*, *9*(1), 381–386.

Michalski, R. S., Carbonell, J. G., & Mitchell, T. M. (2013). *Machine learning: An artificial intelligence approach*. Springer Science & Business Media.

Morella, P., Lambán, M. P., Royo, J. A., & Sánchez, J. C. (2021). The importance of implementing cyber physical systems to acquire real-time data and indicators. *J*, *4*(2), 147–153.

Mosallam, A., Medjaher, K., & Zerhouni, N. (2016). Data-driven prognostic method based on Bayesian approaches for direct remaining useful life prediction. *Journal of Intelligent Manufacturing*, *27*, 1037–1048.

Nabila, E. A., Santoso, S., Muhtadi, Y., & Tjahjono, B. (2021). Artificial intelligence robots and revolutionizing society in terms of technology, innovation, work and power. *IAIC Transactions on Sustainable Digital Innovation (ITSDI)*, *3*(1), 46–52.

Namuduri, S., Narayanan, B. N., Davuluru, V. S. P., Burton, L., & Bhansali, S. (2020). Deep learning methods for sensor based predictive maintenance and future perspectives for electrochemical sensors. *Journal of The Electrochemical Society*, *167*(3), 037552.

Nguyen, H. D., Tran, K. P., Zeng, X., Koehl, L., Castagliola, P., & Bruniaux, P. (2019). Industrial Internet of Things, big data, and artificial intelligence in the smart factory: A survey and perspective. *ISSAT International Conference on Data Science in Business, Finance and Industry*, 72–76.

Onel, M., Kieslich, C. A., Guzman, Y. A., & Pistikopoulos, E. N. (2018). Simultaneous fault detection and identification in continuous processes via nonlinear support vector machine based feature selection. In *Computer Aided Chemical Engineering* (Vol. 44, pp. 2077–2082). Elsevier.

Poór, P., Basl, J., & Zenisek, D. (2019). Predictive Maintenance 4.0 as next evolution step in industrial maintenance development. *2019 International Research Conference on Smart Computing and Systems Engineering (SCSE)*, 245–253.

Poór, P., Ženíšek, D., & Basl, J. (2019). Historical overview of maintenance management strategies: Development from breakdown maintenance to predictive maintenance in accordance with four industrial revolutions. *Proceedings of the International Conference on Industrial Engineering and Operations Management, Pilsen, Czech Republic*, 23–26.

Qiao, G., & Weiss, B. A. (2018). Quick health assessment for industrial robot health degradation and the supporting advanced sensing development. *Journal of Manufacturing Systems*, *48*, 51–59.

Rachidi, A., Talbi, A., & Khatory, A. (2013). The new forms of industrial maintenance: Which impact on the performance of the industrial companies?(Case study). *International Journal of Engineering and Advanced Technology (IJEAT), Exploring Innovation, ISSN*, 2249–8958.

Reddy, Y., Viswanath, P., & Reddy, B. E. (2018). Semi-supervised learning: A brief review. *Int. J. Eng. Technol*, *7*(1.8), 81.

Rousopoulou, V., Nizamis, A., Giugliano, L., Haigh, P., Martins, L., Ioannidis, D., & Tzovaras, D. (2019). Data analytics towards predictive maintenance for industrial ovens: A case study based on data analysis of various sensors data. *Advanced Information Systems Engineering Workshops: CAiSE 2019 International Workshops, Rome, Italy, June 3-7, 2019, Proceedings 31*, 83–94.

Saxena, A., & Saad, A. (2007). Evolving an artificial neural network classifier for condition monitoring of rotating mechanical systems. *Applied Soft Computing*, *7*(1), 441–454.

Scalabrini Sampaio, G., Vallim Filho, A. R. de A., Santos da Silva, L., & Augusto da Silva, L. (2019). Prediction of motor failure time using an artificial neural network. *Sensors*, *19*(19), 4342.

Scarf, P. A. (2007). A framework for condition monitoring and condition based maintenance. *Quality Technology & Quantitative Management*, *4*(2), 301–312.

Sharifi, R., & Ebrahimi, M. (2011). Detection of stator winding faults in induction motors using three-phase current monitoring. *ISA Transactions*, *50*(1), 14–20.

Silvestrin, L. P., Hoogendoorn, M., & Koole, G. (2019). A Comparative Study of State-of-the-Art Machine Learning Algorithms for Predictive Maintenance. *SSCI*, 760–767.

Sullivan, G., Pugh, R., Melendez, A. P., & Hunt, W. D. (2010). *Operations & maintenance best practices-a guide to achieving operational efficiency (release 3)*. Pacific Northwest National Lab.(PNNL), Richland, WA (United States).

Susto, G. A., Schirru, A., Pampuri, S., McLoone, S., & Beghi, A. (2014). Machine learning for predictive maintenance: A multiple classifier approach. *IEEE Transactions on Industrial Informatics*, *11*(3), 812–820.

Vepa, R. (1992). Introduction to Fuzzy Logic and Fuzzy Sets. In *Application of Artificial Intelligence in Process Control* (pp. 146–163). Elsevier.

Vingerhoeds, R. A., Janssens, P., Netten, B. D., & Fernández-Montesinos, M. A. (1995). Enhancing off-line and on-line condition monitoring and fault diagnosis. *Control Engineering Practice*, *3*(11), 1515–1528.

Von Birgelen, A., Buratti, D., Mager, J., & Niggemann, O. (2018). Self-organizing maps for anomaly localization and predictive maintenance in cyber-physical production systems. *Procedia Cirp*, *72*, 480–485.

Wang, P., Youn, B. D., & Hu, C. (2012). A generic probabilistic framework for structural health prognostics and uncertainty management. *Mechanical Systems and Signal Processing*, *28*, 622–637.

What is Deep Learning? | IBM. (2021, October 21). https://www.ibm.com/topics/deep-learning

Xiang, S., Huang, D., & Li, X. (2018). A generalized predictive framework for data driven prognostics and diagnostics using machine logs. *TENCON 2018-2018 IEEE Region 10 Conference*, 0695–0700.

Xu, M., David, J. M., & Kim, S. H. (2018). The fourth industrial revolution: Opportunities and challenges. *International Journal of Financial Research*, *9*(2), 90–95.

Yan, R., Gao, R. X., & Chen, X. (2014). Wavelets for fault diagnosis of rotary machines: A review with applications. *Signal Processing*, *96*, 1–15.

Yang, B.-S., Jeong, S. K., Oh, Y.-M., & Tan, A. C. C. (2004). Case-based reasoning system with Petri nets for induction motor fault diagnosis. *Expert Systems with Applications*, *27*(2), 301–311.

Yung, K.-C., Sun, B., & Jiang, X. (2017). Prognostics-based qualification of high-power white LEDs using Lévy process approach. *Mechanical Systems and Signal Processing*, *82*, 206–216.

Zhan, Y., & Mechefske, C. K. (2007). Robust detection of gearbox deterioration using compromised autoregressive modeling and Kolmogorov–Smirnov test statistic—Part I: Compromised autoregressive modeling with the aid of hypothesis tests and simulation analysis. *Mechanical Systems and Signal Processing*, *21*(5), 1953–1982.

Zhang, Y., Ding, X., Liu, Y., & Griffin, P. J. (1996). An artificial neural network approach to transformer fault diagnosis. *IEEE Transactions on Power Delivery*, *11*(4), 1836–1841.

Zhou, A., Yu, D., & Zhang, W. (2015). A research on intelligent fault diagnosis of wind turbines based on ontology and FMECA. *Advanced Engineering Informatics*, *29*(1), 115–125.