



RADemics

Machine Learning Based Predictive Maintenance Models in Manufacturing and Production Systems

Kannan Gopu, S. Geethanjali, S.
Vellingiri

SREE SASTHA INSTITUTE OF ENGINEERING AND
TECHNOLOGY, SSM INSTITUTE OF ENGINEERING
AND TECHNOLOGY, COIMBATORE INSTITUTE OF
TECHNOLOGY

Machine Learning Based Predictive Maintenance Models in Manufacturing and Production Systems

¹Kannan Gopu, Assistant Professor /Mechanical Engineering , Sree Sastha Institute Of Engineering And Technology, Chembarambakkam, Tamil Nadu - 600124, Email ID: reykan087@gmail.com

²S. Geethanjali, Assistant Professor /EEE, SSM Institute Of Engineering And Technology, Dindigul - 624002, Email ID: geethanjalisubramani11@gmail.com

³S. Vellingiri, Assistant Professor /Mechanical Engineering, Coimbatore Institute of Technology, Coimbatore – 641014, Email ID: vellingiri.s@cit.edu.in

Abstract

The integration of machine learning into predictive maintenance (PdM) frameworks has emerged as a transformative approach for enhancing reliability, reducing downtime, and optimizing operational efficiency in manufacturing and production systems. The effectiveness of data-driven PdM strategies is often constrained by imbalanced datasets, limited failure annotations, and heterogeneous sensor signals that characterize real-world industrial environments. This chapter presents a comprehensive examination of hybrid machine learning architectures designed to address the limitations posed by sparse and imbalanced data conditions. Emphasis is placed on the fusion of rule-based logic with statistical learning and deep neural models to achieve robust, interpretable, and scalable maintenance intelligence. The chapter explores advanced feature engineering techniques, sensor data fusion strategies, and data augmentation mechanisms to enhance predictive accuracy and generalizability. Visualization methods for explainable decision support are discussed to bridge the gap between model outputs and actionable maintenance insights. Real-world constraints such as computational efficiency, latency, and domain adaptability are considered throughout the chapter, offering a holistic view of deploying hybrid learning frameworks in practical industrial settings. The proposed methodologies not only address current challenges in PdM model development but also lay the groundwork for next-generation, adaptive maintenance solutions in smart manufacturing ecosystems.

Keywords: Predictive Maintenance, Hybrid Machine Learning, Imbalanced Data, Sensor Fusion, Explainable AI, Industrial Intelligence

Introduction

The industrial sector is undergoing a paradigm shift with the adoption of smart manufacturing systems driven by Industry 4.0 principles [1]. In this context, ensuring the continuous operation and reliability of machinery has become increasingly critical to maintaining productivity, minimizing costs, and sustaining competitive advantage [2]. Predictive maintenance (PdM) has

emerged as a cornerstone of this transformation, enabling early fault detection and proactive intervention through data-driven insights [3]. Unlike reactive or time-based maintenance, PdM relies on the continuous monitoring and analysis of machine data to forecast failures before they occur [4]. This transition from scheduled interventions to intelligent predictions marks a significant advancement in operational efficiency and asset management. Realizing the full potential of predictive maintenance in industrial settings is often hindered by challenges related to data quality, availability, and interpretability [5].

Real-world manufacturing environments typically involve heterogeneous machines, variable operational conditions, and non-uniform data acquisition systems [6]. These factors result in datasets that are often sparse, noisy, and heavily imbalanced—where fault occurrences represent a minor fraction of total observations [7]. Most conventional machine learning models, which assume data abundance and balanced class distributions, fail to perform effectively under such constraints [8]. Models trained on skewed datasets may achieve high overall accuracy while completely neglecting rare but critical failure instances [9]. The absence of adequate failure data, coupled with frequent missing values and inconsistent labeling, exacerbates the difficulty of building robust and generalizable predictive systems. These limitations highlight the inadequacy of single-model approaches and underscore the necessity for more adaptive, integrative frameworks capable of handling the realities of industrial data [10].

To address these constraints, hybrid machine learning architectures are gaining traction as a viable solution [11]. These architectures combine the strengths of different computational models—ranging from rule-based systems and statistical methods to deep learning networks—to enhance both predictive accuracy and system robustness [12]. Hybrid approaches can incorporate domain knowledge through explicit rules, process historical trends using statistical models, and capture nonlinear relationships with deep learning techniques [13]. Such multi-layered systems enable more nuanced interpretation of sensor data and support decision-making even when data is limited or partially observed [14]. By integrating models with complementary strengths, hybrid architectures can address limitations inherent in individual methods, such as the black-box nature of deep learning or the rigidity of rule-based logic. This synergy not only improves performance under constrained data conditions but also enhances model interpretability and operational transparency [15].