ISSN: 2584-0231(Online)
International Journal of Multidisciplinary Research in Arts, Science and Technology
© IJMRAST | Vol. 3 | Issue 5 | May 2025
Available online at: https://ijmrast.com
DOI: https://doi.org/10.61778/ijmrast.v3i5.131

MACHINE LEARNING IN PRODUCTION ENGINEERING: A COMPREHENSIVE REVIEW

Parankush Koul[1]*, Yashavanth Siddaramu[2]

[1,2]Department of Mechanical and Aerospace Engineering, Illinois Institute of Technology, Chicago, Illinois, United States of America - 60616

*Corresponding Author Email Id: pkoul2@hawk.iit.edu

ABSTRACT

The study examines how machine learning (ML) methods can be incorporated into production engineering practices. The paper highlights data preprocessing and cleaning as essential steps to maintain data quality and reliability for ML applications. The review shows the production environment challenges that include missing data values and the presence of outliers along with data inconsistencies. The text explains how advanced automation techniques decrease human involvement while improving feature extraction methods, which produce uniform features across different manufacturing systems. The paper emphasizes that effective model deployment relies on rigorous data engineering pipelines that perform comprehensive data ingestion, transformation, and feature engineering. The review intends to explore the existing ML applications within production engineering while identifying key practices that enable model readiness and reliability.

Keywords: *Machine Learning, Data Partitioning, Feature Engineering, Model Selection, Explainable AI*

INTRODUCTION

ML, which belongs to the artificial intelligence (AI) family, represents a groundbreaking technology that permits computer programs to learn independently from data through a self-learning process, which has received substantial progress, especially with deep learning (DL) applications across different sectors [1]. ML integration in production engineering provides engineers with an all-encompassing framework to implement ML-based models throughout essential production tasks, thereby boosting operational efficiency and lowering costs while promoting sustainable development [2]. This technology transforms manufacturing processes into interconnected systems that support dynamic operations with extensive data capabilities, which define Industry 4.0 [3]. The substantial growth of accessible data throughout recent decades has positioned ML as an effective approach to solve advanced manufacturing problems [4].

ML for production engineering enables computer systems to learn independently from human assistance and modify their actions, which leads to more intelligent and adaptive production processes [4]. Numerous manufacturing sectors utilize this technology because traditional methods fail to manage large data volumes efficiently [5]. Specialized ML approaches need to fulfill production engineering requirements in accuracy, scalability, complexity management, verifiability, and usability, unlike traditional methods, which fall short [6]. Through systematic ML implementation, production optimization and control alongside turndown management and maintenance and intervention planning will benefit substantially because they represent key areas of ML value delivery [2].

Production engineering has witnessed a dramatic increase in the implementation of ML to solve various challenges such as predictive maintenance alongside process optimization and quality control while managing resource scheduling and anomaly detection [4]. ML applications use their data analysis capabilities to detect inefficiencies and predict problems while improving manufacturing processes throughout the production value chain [3]. Predictive maintenance made possible through ML technology is now vital for manufacturing sectors to achieve optimal performance and safety through the analysis of machine data and early identification of anomalies that might disrupt operations [7]. ML integration into production infrastructure organization substantially boosts profitability, especially when new business process integration technologies evolve alongside industrial automation advances [8].

The application of ML within production engineering shows great potential yet encounters numerous obstacles such as data privacy issues, insufficient skilled workforce availability, and essential partnerships between manufacturers and technology vendors [9]. The prevailing limitations in this context require structured ML deployment methods that encompass detailed protocols for data acquisition and preprocessing together with model training procedures [7]. The move toward Industry 4.0 standards in production systems brings substantial potential through ML but requires overcoming intricate integration obstacles to unlock its full advantages [1].

This extensive evaluation examines the dynamic domain of ML within production engineering by analyzing its applications together with its benefits and challenges and projecting its future directions. The study provides a structured analysis of existing research papers and technological developments while showcasing practical applications that illustrate the transformation of production engineering methods through ML across different manufacturing industries. The review integrates knowledge from multiple fields to equip production engineers, researchers, and industrial professionals with comprehensive insights about how ML can transform manufacturing processes to boost efficiency and drive innovation.

STEPS TO IMPLEMENT ML IN PRODUCTION ENGINEERING

A structured end-to-end approach that includes problem identification, data handling, model building, and deployment, as well as continuous monitoring and retraining, is necessary for ML implementation in production engineering [10].

Step 1: Problem Definition and Requirements Gathering

The initial step of this process requires a detailed identification of the production engineering problem that needs resolution [10]. To understand operational goals alongside constraints and performance metrics requires involving relevant stakeholders, specifically for business targets like fault detection accuracy and cost reduction objectives [2]. Break down the problem into specific, measurable objectives that align with data-driven solution methods [11].

Step 2: Data Collection and Organization

High-quality relevant data forms the base for any successful ML initiative [12]. Production engineering data collection encompasses sensor readings along with equipment maintenance logs, process control metrics, production yields, and petrophysical properties, according to sources [6, 2, 12]. The best practices require organizations to perform detailed reviews of available data sources while cataloging existing models and simultaneously evaluating data accuracy and completeness as well as understanding dependencies between data types [2, 12, 13]. The implementation of automated data collection processes, especially for expansive or decentralized datasets, reduces human errors and enhances data quality [14]. Domain expertise input combined with an organized data inventory helps to identify relevant engineering task features effectively and efficiently [15].

Step 3: Data Preprocessing and Cleaning

Production environment data acquisition usually leads to the accumulation of information that includes missing entries and outliers as well as inconsistent records and various data formats [12, 16]. Preprocessing involves several steps: The data preprocessing process includes cleaning data through missing value and outlier handling while also normalizing data or performing feature scaling to validate data consistency [16, 17]. Advanced automation decreases human involvement while improving feature extraction and implementing strong quality control to deliver organized, consistent features throughout various production settings [12]. The development of robust data engineering pipelines through complete data ingestion and transformation processes together with feature engineering is essential for maintaining model readiness and reliability [17, 18].

Step 4: Feature Engineering and Selection

The effectiveness of ML models in production engineering relies heavily on proper feature extraction and engineering, which demands extensive knowledge of both the data characteristics and engineering principles [19]. In this stage, new features are extracted from raw data through variable transformations and domain knowledge integration to generate informative inputs for ML models [19, 20]. The process of selecting features from complex datasets is accelerated by automated systems, which help extract petrophysical properties for reservoir simulation and analyze pump conditions using sensor data [12, 19].

Step 5: Data Partitioning

In order to develop effective models, the dataset must be divided into training, validation, and test sets [21]. The approach provides unbiased model evaluation while supporting hyperparameter optimization and performance benchmarking [22]. The partitioning process must maintain representative distributions of operational scenarios found in production engineering settings [19].

Step 6: Model Selection

The choice of a suitable ML algorithm requires consideration of the problem type and dataset characteristics as well as desired performance attributes [22, 23]. Random forests (RFs) perform better than many other algorithms because they handle diverse data effectively [22, 24]. When working with time-series prediction or process optimization tasks along with failure detection, it is important to evaluate tree-based models together with neural networks and support vector machines (SVM) as well as hybrid methods [20, 25]. The selection of the final model requires comprehensive experimentation with various algorithms followed by performance comparisons [26].

Step 7: Model Training and Hyperparameter Tuning

The process of model training applies the chosen algorithm to cleaned training data using iterative optimization approaches [27]. Grid search and automated tools applied during hyperparameter tuning improve overall model output [25]. Domain-relevant input variability along with feature engineering leads to models that are both more interpretable and more effective [19].

Step 8: Model Validation and Testing

Validation and testing processes verify that the trained ML model maintains its performance on new data while fulfilling practical deployment standards [28]. Implement complete testing approaches that encompass property-based testing along with quality attribute scenarios and risk assessment for deployment alongside continuous validation frameworks according to references [29, 30, 31]. Robust test cases integrate operational constraints with integration requirements to prevent deployment failures [32]. The development process emphasizes reproducibility and interpretability to build stakeholder trust while meeting regulatory standards [33].

Step 9: Deployment into Production Environment

The model deployment process includes embedding the validated model into an active production engineering system [10]. Implement containerization tools such as Docker and Kubernetes to achieve portable deployments while maintaining scalable operations and resource isolation [34]. ML model deployment benefits from ML Operations (MLOps) frameworks such as Kubeflow and MLflow, which provide lifecycle management alongside versioning with automated scaling and system health monitoring capabilities [35]. Automated deployment pipelines provide consistent maintenance for production system integration [36].

Step 10: Continuous Monitoring and Performance Tracking

ML models deployed in production need ongoing monitoring to ensure performance metrics and prediction accuracy remain stable while guaranteeing system reliability [37, 38]. Use drift detection methods to identify changes in data distributions, such as data drift and concept drift, while creating automated alerts to report performance drops [39, 40]. Organizations utilize diagnostic and logging tools to conduct root cause analysis while ensuring they meet operational and regulatory requirements [39].

Step 11: Feedback Loop and Model Retraining

Establish automated feedback loops for retraining ML models to maintain their accuracy and relevance through production data changes and evolving requirements [41, 42]. The frequency of retraining models, whether periodic or event-driven, must consider operational expenses, while probabilistic models or predictive methods help find the best intervals for retraining [43, 44]. Continuous model improvement and system resilience depend on integrating real-time monitoring with data versioning and retraining pipelines [42, 45]. Self-updating systems in highly dynamic environments benefit from reinforcement learning (RL) or adaptive strategies [46].

Step 12: Documentation, Collaboration, and Maintenance

Ensure complete documentation of modeling activities together with data management and deployment processes throughout the entire operation to support reproducibility and troubleshooting while facilitating knowledge transfer [33]. Ensure data scientists, engineers, operators, and domain experts work together so all decisions stay relevant to context and business goals. Strong governance procedures combined with established institutional processes ensure sustainable implementation of models throughout their life cycle.

Table 1: Steps for Implementing ML in Production Engineering—Summary

Step	***Key Activities***
1. Problem Definition	Define objectives, engage stakeholders, set metrics
2. Data Collection & Organization	Review sources, automate collection, inventory features, validate quality
3. Data Preprocessing & Cleaning	Handle missing/outlier/noise, normalize/scale, automate QC
4. Feature Engineering & Selection	Derive features, integrate domain knowledge, automate extraction
5. Data Partitioning	Split data into train/validation/test sets, maintain operational representativeness
6. Model Selection	Select algorithms (e.g., RFs, neural nets), evaluate multiple approaches
7. Model Training & Tuning	Fit model, optimize hyperparameters, iterative improvement
8. Model Validation & Testing	Comprehensive testing (QA scenarios, property based), risk assessment
9. Deployment	Integrate with production, use containers, adopt MLOps for automation and scaling
10. Continuous Monitoring	Track accuracy, detect drift, establish alerts, root-cause diagnostics

Step	Key Activities
11. Feedback Loop & Retraining	Automated retraining pipelines, optimize schedules, incorporate new data
12. Documentation & Collaboration	Maintain documentation, foster cross-functional teamwork, apply governance

BENEFITS AND APPLICATIONS OF ML IN PRODUCTION ENGINEERING

Benefits of ML in Production Engineering

Production engineering achieves improved productivity and manufacturing resilience while becoming more cost-effective through the diverse benefits provided by ML.

- *Efficiency Improvement and Cost Reduction*: ML delivers significant operational efficiency improvements by automating standard processes while optimizing production timelines and adapting operational parameters through real-time data analysis [48]. ML models help to detect production bottlenecks while minimizing waste and refining manufacturing processes, which leads to decreased energy usage together with lower labor expenses and enhanced throughput [49]. ML-driven automated defect detection speeds up quality checks while minimizing expenses from manual inspections and faulty product reprocessing [50]. Research shows that the use of ML in industrial control systems generates cost savings between 10–20% alongside energy reductions ranging from 1–10% [49].
- *Predictive Maintenance and Reliability*: The standard approach to maintenance uses scheduled or reactive interventions, which create unnecessary downtime and cause unexpected equipment breakdowns [51]. ML-driven predictive maintenance uses sensor inputs alongside historical information to predict when equipment will fail, which enables preventative actions that boost both uptime and the lifespan of machinery [52]. Applications with model accuracy rates above 90% in real-world textile manufacturing settings successfully reduced unplanned downtime and maintenance expenses, according to research [52].
- *Quality Improvement and Consistency*: ML techniques boost product quality through precise identification of key process variables that cause defects, which allows continuous process optimization [53]. The integration of feature selection algorithms with ML techniques enables the determination of ideal parameter settings for defect-free production, demonstrated by more than 92% accuracy in medical mask manufacturing [53]. Quality control systems that rely on ML demonstrate superior capabilities in identifying minor anomalies within production data, which aids organizations in achieving zero-defect manufacturing standards [54].
- *Enhanced Decision-Making*: The use of ML to discover production data patterns enables decision-makers to operate more intelligently and swiftly despite unpredictable and variable conditions [48]. The ability to quickly adapt production processes remains essential for industries involved in process manufacturing and discrete part assembly, which face shifting production needs [13].
- *Automation and Flexible Manufacturing*: ML-driven automation and robotics have transformed manufacturing in the context of Industry 4.0 [55]. ML technologies provide adaptive control capabilities and enable quick production line adjustments and cooperative human-robot operations that improve manufacturing scalability and process flexibility [56].
- *Sustainability and Waste Reduction*: Organizations now widely use ML models to enhance resource usage efficiency while minimizing energy consumption and waste generation [57]. Advanced systems such as

EcoEfficientNet process production data instantaneously to detect inefficiencies and generate data-driven improvement suggestions, which resulted in waste reduction by up to 30% during testing [57].

- *Supply Chain Optimization*: ML-enabled predictions enable complex global supply chains to optimize demand forecasts and logistics management while reducing lead times and inventory requirements [58]. Production engineers who use algorithms such as regression, clustering, and networks achieve better forecasts while reducing costs and enhancing supplier performance and delivery reliability [59].

Applications of ML in Production Engineering

The best way to understand ML's impact on production engineering lies in studying its fundamental application areas.

- *Predictive Maintenance*: ML achieves one of its key applications in industrial settings through predictive maintenance practices [7]. Operational and sensor data streams are analyzed by supervised and unsupervised learning algorithms along with RL to predict machinery failures and suggest maintenance schedules [60]. Research in the oil and gas industry, together with automotive manufacturing and textile production, shows substantial decreases in unplanned downtime and maintenance expenses plus waste reduction while equipment reliability improved [19, 61].

  Notable achievements include:

  - Automated alert systems in real-time fault detection frameworks serve to notify operators of operational anomalies [62].
  - ML techniques enable better maintenance scheduling, which leads to longer machine durability and lower labor costs [7, 51].
  - Internet of Things (IoT) infrastructure integration enables seamless monitoring of conditions and data collection capabilities [55].

- *Process Optimization*: ML demonstrates superior capability in detecting inefficiencies and refining manufacturing processes [2]. ML models process large datasets to assess how operating conditions and raw material differences impact product quality and throughput performance [55]. Real-time dynamic process adjustments in Industry 4.0 manufacturing settings lead to higher production yields alongside reduced downtime and faster responses to process variability [3].

  For example:

  - ML applications to optimize parameters in grinding and classification operations led to better quality stability and enhanced overall process dependability [63].
  - The process control accuracy of simulation-driven optimization frameworks increases by integrating ML with physics-based models [2].

- *Quality Control and Defect Detection*: ML transforms industrial quality assurance by:

  - ML technology enables automated defect detection and classification by enhancing image recognition and sensor analysis with predictive modeling techniques [54].
  - Adaptive quality management systems drive the process by adjusting parameters according to detected anomalies, which ensures consistently high product quality [49].
  - Production data enables the implementation of "zero-defect" strategies through continuous learning practices [64].

The latest research demonstrates how ML achieves precise detection of defective products while offering process enhancement suggestions and delivering practical feedback to operators [53, 65].

- *Automation and Robotics*: Modern production automation relies heavily on the integration of ML with robotics.
    - ML-enabled collaborative robots (cobots) ensure human safety through adaptive task management and workflow optimization [56, 66].
    - ML-enhanced automation provides operations with increased speed and precision while maintaining flexibility when dealing with complex or unpredictable environments [67, 68].
    - Significant advancements in real-time control and vision, as well as grasping and adaptive motion planning, enable new opportunities for customized manufacturing growth and production scaling [69].
- *Supply Chain Management and Logistics*: In supply chain management, ML's predictive power exhibits substantial potential.
    - ML improves demand forecasting while optimizing inventory alongside supplier prediction, which helps reduce both under- and over-stock occurrences [59].
    - Causal ML enables "what-if" analysis, which aids risk mitigation and improves resilience against disruptions [70].
    - The global production cost efficiency improves through automated selection of optimal manufacturing and delivery locations [58].
- *Sustainability and Resource Optimization*: ML contributes to sustainable manufacturing by:
    - The use of continuous data-driven monitoring and adjustments enables optimal energy and material usage [57].
    - Circular economy practices benefit from ML support, which focuses on waste reduction and recycling plus eco-friendly operational methods [57].
    - Real-time emissions monitoring together with adaptive process control helps achieve better environmental compliance [49].

PAST PROJECTS IN COMPANIES IMPLEMENTING ML IN PRODUCTION ENGINEERING

- *IBM*: IBM is adopting MLOps to centralize ML model management and deployment in production. With this automation of many processes in the ML lifecycle, IBM tries to make the application scalable and more efficient. MLOps helps make sure that ML models can scale, adapt, and learn from data changes while delivering high performance and accuracy. Data scientists and IT teams are closely integrated with the company, and the work culture is very robust; models are always getting improved through feedback loops [71].
- *Google*: Google is the first company to adopt ML in production engineering with MLOps. It works on the application of ML in business applications and on the challenges of running ML models. By acquiring infrastructure for continuous integration and deployment, Google has made model development-to-deployment workflows much smoother, and models remain scalable. That way Google can effectively take care of the lifecycle of ML models and solve problems in the real world [72].

- *Amazon*: Amazon Robotics department used Amazon SageMaker to create a powerful ML model to automate inventory control in its fulfillment centers. The intent detection system developed as part of this work is based on DL techniques that are conditioned on large video sets to detect stowing behaviors without manual scanning. Amazon was able to make huge use of the resources and processing power provided by AWS services, which led to reduced cost and increased functionality [73].
- *Microsoft*: Microsoft has prioritized helping teams build AI applications with ML workflow best practices. Such as applying MLOps principles to support cross-team coordination between traditional software and ML teams for scalable and reliable ML systems [74, 75]. The learnings from software teams' analysis have helped to build models that facilitate model training and deployment and, in the end, have helped bring ML into production engineering.
- *Netflix*: ML is used at Netflix in all areas of operations, which makes the business applications, such as content recommendation and media knowledge, very helpful. Netflix's Metaflow framework is used to create a strong ML platform on which to deploy and manage various ML models in the organization. This makes data and model building more efficient for teams working on it, which enables better predictive abilities and content delivery efficiencies [76].

OVERVIEW OF PAST RESEARCH ON ML IN PRODUCTION ENGINEERING

The number of articles covered in this review on the ML in Production Engineering are shown in Figure 1 from 2019 through 2024.

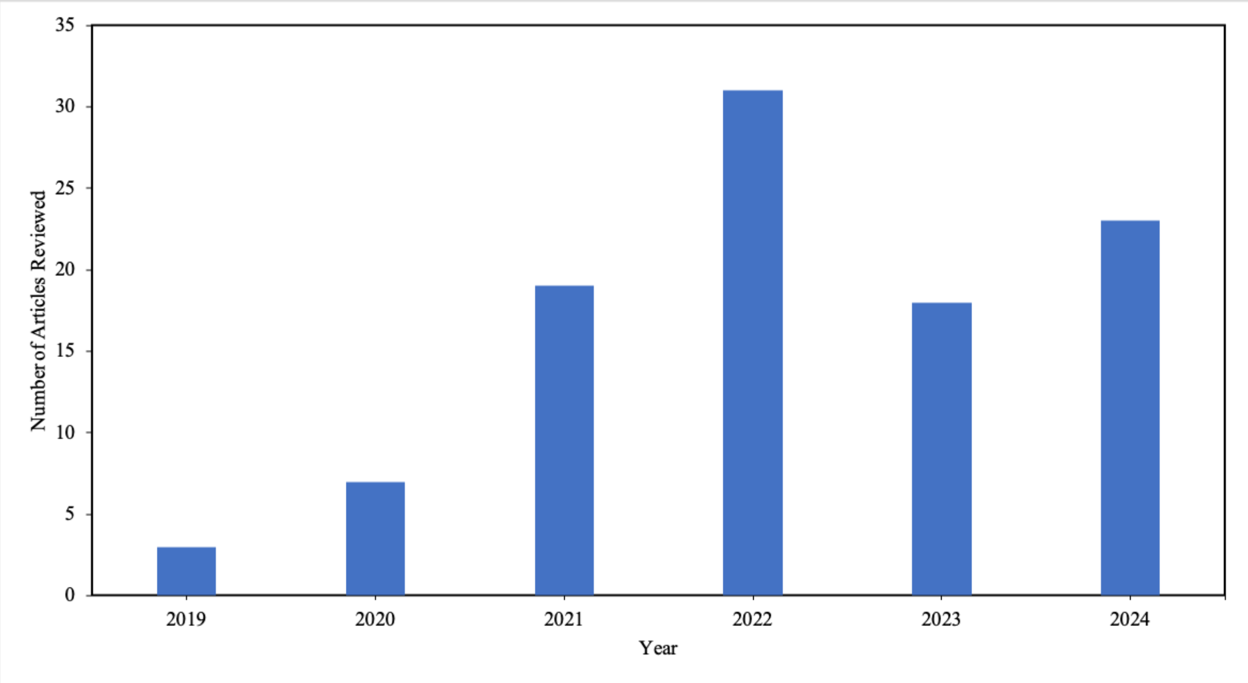

Fig. 1: Number of articles on ML in Production Engineering vs. Year

Table 2 below shows a quantitative distribution by publisher of the number of articles related to the applications of ML in Production Engineering.

Table 2: Number of articles from different publishers reviewed on the applications of ML in Production Engineering

Publisher	Number of Articles Reviewed
Elsevier	29
MDPI	28
Springer	18
IEEE	5
ACM Digital Library	4
Taylor & Francis	3
ACS Publications	1
ALL ACADEMIC RESEARCH	1
arXiv (Cornell University)	1
Association for Computational Linguistics (ACL)	1
CILAMCE	1
Darcy & Roy Press	1
elibrary.ru	1
Frontiers	1
Opast Publishing Group	1
PNT SCIENTIFIC TEMPER PRIVATE LIMITED	1
SPE International	1
Wiley	1
Yandy Scientific Press	1
Zeus Press	1
Total	**101**

Colyer (2019) demonstrated that data validation and software engineering are essential for integrating ML into production systems by utilizing Google's extensive data validation strategies and Microsoft's methods for incorporating ML into product ecosystems. Model performance degrades over time if changes in data distributions occur without adequate monitoring and lifecycle management [28]. Soto et al. (2019) introduced an online ML framework specifically designed to detect product failures early in Industry 4.0 environments. The combination of RFs and gradient boosting with a discrete-event simulation model for surface-mount technology (SMT) lines allowed real-time failure detection, which enhanced equipment performance and minimized waste [77]. Paul et al. (2019) introduced an iterative ML model that predicts temperature patterns in additive manufacturing by using extremely randomized trees to maintain prediction accuracy under 1% mean absolute percentage error (MAPE) and enable real-time process management [78].

Rodríguez et al. (2020) designed an intelligent decision support system using fuzzy logic and regression trees to improve decision-making processes for closed-loop supply chains in hospital laundry operations. The developed system produced classification prediction accuracy exceeding 75% and outsourcing decision accuracy above 80%, which matched expert evaluations closely [79]. Dučić et al. (2020) used ML models, including neural networks and support vector regression (SVR), to predict alloying additives for white cast iron melting processes to enhance process control. The research results showed that neural networks achieved better prediction accuracy than SVR, which indicates their practical use for chemical composition optimization [80]. San-Payo et al. (2020) investigated quality control methods for clothing manufacturing through a system that combined a convolutional neural network (CNN) with InceptionResNET architecture and an adapted Mondrian forest for incremental learning. Their technology provided defect identification and adaptive classification capabilities for clothing components through images captured by mobile devices and delivered dependable results across multiple evaluation standards [81].

In their 2020 study on Laser Powder Bed Fusion (LPBF) for additive manufacturing, Gaikwad et al. developed the Sequential Decision Analysis Neural Network (SeDANN) model to process heterogeneous pyrometer and high-speed video data. The model achieved better results than conventional black-box methods because it included physical process knowledge, which led to substantial accuracy gains in track height deviation predictions for in-situ quality assurance [82]. Kaya et al. (2020) developed a Sensor Failure Tolerable model that utilizes a Single Plurality Voting System (SPVS) to improve sensor reliability in food production. The integration of K-Nearest Neighbor (kNN) and Decision Trees classifiers within an SPVS ensemble enabled consistent high accuracy in beef quality predictions despite sensor failures, which validated system robustness during imperfect conditions [83]. Orru and colleagues (2020) introduced ML methods to predict faults in centrifugal pumps utilized within the oil and gas sector. Their research demonstrated that SVM and Multilayer Perceptron (MLP) used on temperature, pressure, and vibration sensor data achieved reliable early fault detection, which led to enhanced predictive maintenance and operational efficiency [84].

The research by Arellano-Espitia et al. (2020) introduced a deep-learning fault diagnosis method for electromechanical systems that combines unsupervised stacked auto-encoders with supervised discriminant analysis to handle the complexity and multi-fault conditions present in modern manufacturing settings. Results indicated that the methodology achieved enhanced adaptability and application simplicity, proving its effectiveness across different system states while enabling more efficient maintenance procedures [85]. Lutz et al. (2021) developed an in-situ material batch identification system for machining operations by utilizing supervised ML techniques, including SVM, RFs, and Artificial Neural Networks (ANNs) to analyze torque signals. The SVM model delivered the highest prediction accuracy in their five material batch experiments and significantly surpassed logistic regression, which allowed for real-time optimization of production processes based on material characteristics [86]. Khayyati and Tan (2021) introduced a structured ML method to develop production control policies based on insights from vast production datasets to adjust control decisions effectively. The approach improved production efficiency through waste reduction and better resource allocation while demonstrating how data-driven strategies effectively support dynamic manufacturing settings [87].

Researchers Sauter and team in 2021 developed a method to detect grinding burn during production by using ML techniques to evaluate sensor readings from both acoustic emission and power channels. The research demonstrated exceptional classification accuracy of nearly 99% for grinding burn detection, which showed that multi-sensor data in real-time monitoring performs much better than conventional post-process techniques [88]. Rai et al. (2021) examined ML's integration into the Industry 4.0 framework to show its transformative effects on manufacturing through predictive maintenance and quality control as well as data-based decision-making. The research showed that smart sensors and devices are crucial for real-time data analysis while recognizing obstacles such as data privacy concerns and system integration difficulties [89]. Simeone et al. (2021) created a cloud manufacturing framework that uses ML to provide manufacturing solutions through analysis of historical customer data. The neural network-based recommendation system enhanced customer satisfaction through precise solution predictions, which showed a decrease in Root Mean Square Error (RMSE) values [90].

Through their empirical analysis of 3,000 Google production ML pipelines, Xin et al. (2021) discovered that traditional ML research emphasizes model training yet shows that 80% of computational time is dedicated to data management and validation tasks. The researchers identified inefficiencies within the pipelines while recommending ML-based optimization methods to minimize computational waste and enhance data preparation efficiency [91]. Kang and their team (2021) investigated production line predictive maintenance by applying MLP neural networks to estimate equipment Remaining Useful Life (RUL). ML effectiveness in maintenance strategy enhancement was shown through its capability to detect equipment failures early, which resulted in better resource management and operational scheduling [92]. Burggräf et al. (2021) tackled the problem of forecasting assembly start delays within manufacturing settings. A total of 24 prediction models were developed using classification and regression methods, which revealed that the gradient boosting classifier excelled at predicting missing components and assembly delays. The study highlighted how making accurate predictions at the right time helps achieve delivery deadlines and boosts manufacturing efficiency [93].

The research by Bampoula and his team in 2021 targeted predictive maintenance for cyber-physical production systems through the use of Long Short-Term Memory (LSTM) autoencoders. The study showed that LSTM autoencoders excel at classifying data from machines and sensors, which enables companies to move from scheduled maintenance to predictive maintenance models that function in real-time. This method enhanced operational efficiency through machinery life predictions while simultaneously minimizing system downtime [94]. Benbarrad et al. (2021) created an intelligent machine vision model tailored for quality control purposes in manufacturing settings. The researchers implemented ML techniques to identify defective products while predicting the best production parameters and achieved improvements in product quality alongside process efficiency. Their comparative study showed that the machine vision model functioned as a reliable defect detector while offering valuable process improvement insights within Industry 4.0 [95]. Researchers Siler et al. (2021) utilized unsupervised ML algorithms by implementing non-negative matrix factorization with k-means clustering (NMFk) to determine geothermal production drivers within Nevada's Brady geothermal field. The study showed that macro-scale faults and fault step-overs significantly influence geothermal production through an analysis of 3D geologic data, which also enhanced understanding of subsurface processes [96].

Szajna and colleagues (2021) developed the Wire Label Reader (WLR) device, which uses deep CNNs (DCNNs) to identify wire markings during industrial assembly processes. Their study demonstrated effective DCNN deployment, reaching a remarkable 99.7% accuracy rate for wire marking detection. The introduction of this technology reduced both assembly time and energy consumption, which led to improvements in operational efficiency as well as sustainability in production settings [97]. Wang et al. (2021) conducted research on offshore field reservoir characterization using an ensemble RF ML algorithm. The model processed field production and injection data to forecast oil saturation profiles and attained above 90% coefficient of determination (R^2) values with less than 6% RMSE. The study proved that ML could improve reservoir management through existing field data while providing a reliable substitute for traditional approaches that depend on expensive and limited data sources [98]. Wang and his team (2021) utilized ML techniques like Gaussian Process Regression (GPR) and CNN models to enhance shale gas production optimization. The researchers found that GPR models achieved superior performance compared to other ML approaches by obtaining a R^2 of 99.99% and an RMSE of 0.0127. The application of ML to well parameter optimization led to a considerable Net Present Value (NPV) increase while demonstrating ML's capability to expedite optimization processes and enhance well design decision-making [99].

The researchers Leon-Medina et al. (2021) created a multivariate time series DL model based on Gated Recurrent Units (GRUs) to estimate temperatures in ferronickel production electric arc furnace (EAF) linings. The research model achieved high accuracy while maintaining an RMSE measurement of 1.19°C and demonstrated the significance of data preprocessing and structural health monitoring within extreme industrial conditions [100]. Yoo and Kang (2021) implemented an explainable AI framework for estimating manufacturing costs based on 3D computer-aided design (CAD) models. The team developed a DL framework incorporating 3D gradient-weighted class activation mapping to display machining features and their related costs, which supports data-driven design choices and immediate cost feedback in the initial design stages [101]. The research by Moretti et al. (2021) solved the assembly line feeding challenge by combining supervised ML techniques together with optimization models to develop effective part-feeding strategies. The truck manufacturing case validated their hybrid method, which showed that the ML model achieved 83.1% of optimal assignments while resulting in a slight 6.1% cost increase [102].

In their 2021 study, Griffin et al. demonstrated that supervised learning techniques such as Classification and Regression Trees (CART) and neural networks achieved up to 97% accuracy in identifying galling wear in sheet metal stamping tools through acoustic emission-based monitoring, significantly surpassing unsupervised methods [103]. Badarinath et al. (2021) created ML substitutes for finite element analysis that estimate one-dimensional beam stress distributions through ANNs, which deliver superior accuracy and enable real-time maintenance planning [104]. Stauder and Kühl (2022) tackled sequence management in mixed-model automotive assembly lines by applying supervised learning to product-related features to predict sequence deviations, which resulted in a 44% detection rate for delayed vehicles with 50% precision while integrating a continuous learning module to adapt model performance [105].

Garouani et al. (2022) introduced a meta-learning-based tool called AMLBID, which automates ML algorithm selection and configuration for industrial big data while addressing manufacturing engineers' lack of ML expertise. The tool demonstrated superior predictive accuracy and computational efficiency over established AutoML software TPOT and Auto-sklearn when tested with 400 real-world datasets and reached near-constant time complexity [106]. Tan and Nhat (2022) examined thermoforming composite materials using ANNs with image analysis from finite element simulations to optimize process parameters and prediction quality. The implementation of their method achieved a 17.4% slip-path length reduction and expanded optimal shear angle regions by 31%, demonstrating how ANNs effectively minimize defects and improve formability [107]. Zhang et al. (2022) utilized SVR along with grey relation analysis (GRA) to enhance the optimization of gluing parameters during particleboard production. The researchers created a model that achieved high internal bond strength prediction accuracy through 724 training and 181 testing datasets, thus showing real-time quality control capabilities in industrial environments [108].

In 2022, Grishina and Sorokin improved Spoken Language Understanding (SLU) models in production systems through the Local-to-Global Learning (LGL) method, which resulted in a 7.3% decrease in error rates and training time reductions by up to 25%. The research focused on how incremental learning can adapt to real-world limitations of changing and restricted data availability [109]. Salem et al. (2022) built a predictive ML model to identify well integrity failures in gas-lift artificial lift systems using ten years of field data from 800 wells. The researchers achieved perfect prediction accuracy by testing 11 ML algorithms and discovered that extreme and categorical boosting models performed best, while their application of domain-specific evaluation metrics alongside physical risk equations created a strong and interpretable tool for optimizing safety and resources in oil and gas operations [110]. Hojati et al. (2022) implemented time-series imaging and CNN technology to track and forecast machining conditions when slot milling titanium alloys. Their model attained high precision (88%) and Area Under the Curve (AUC) scores above 0.95 by transforming real-time Computer numerical control (CNC) machine signals into Gramian Angular Field (GAF) images and implementing data augmentation to balance the dataset, which helped to identify important machining scenarios like tool wear and breakage [111].

Natanael and Sutanto (2022) demonstrated the use of low-cost components like Arduino together with basic sensors for a tube-filling machine's predictive maintenance system based on RF regression, achieving 88% accuracy while increasing Overall Equipment Effectiveness (OEE) by 13.10% and decreasing unplanned failures by 62.38% [112]. Tama et al. (2022) investigated malfunction detection through acoustic signal analysis by building an EfficientNet-based weighted ensemble DL system that surpassed individual classifiers and achieved AUC scores of 0.9981 for pumps and 0.9993 for valves when tested with the Malfunctioning Industrial Machine Investigation and Inspection (MIMII) dataset [113]. The 2022 study by Cao et al. introduced a hybrid AI methodology in their Knowledge-based System for Predictive Maintenance in Industry 4.0 (KSPMI) system, which combined ML with symbolic AI through domain ontologies and Semantic Web Rule Language (SWRL) rules alongside chronicle mining to process diverse industrial datasets and achieved superior predictive accuracy and adaptability for both real and synthetic data scenarios [114].

The research by Motz et al. (2022) evaluated different hyperparameter optimization methods to assess their effectiveness for production environments and emphasized how various application needs complicate the process of selecting appropriate techniques. The structured approach to benchmarking combined with empirical evidence showed that making informed decisions about hyperparameter optimization leads to improved ML performance in production settings [115]. Feng and Li (2022) developed an RL-based decision model for predictive maintenance in multistage production systems that integrated system dynamics and bottleneck identification. The proposed model achieved better results than traditional maintenance approaches through cost reduction and throughput enhancement, which demonstrated the significance of dynamic maintenance decisions based on data analytics [116]. Bhowmik (2022) explored the complete lifecycle of ML deployment in production environments through an analysis of strong ML pipeline architecture. His research highlighted practical difficulties such as data drift and model decay along with the need for ongoing monitoring and automated processes throughout data validation to deployment [117].

Rom et al. (2022) demonstrated the application of advanced ML methods such as Residual Neural Networks (ResNets) within production engineering to predict both cutting forces during machining and particle behavior during plasma spraying. The research combined mathematical modeling techniques with ML through the use of ensemble Kalman filtering (EnKF) to improve prediction accuracy while decreasing computational demands [118]. Bulaev (2022) developed an extensive framework for deploying ML models in production settings, including key steps like goal setting and model training, but emphasized implementation guidelines over algorithm design [119]. Song et al. (2022) conducted a comparative study of ML models like Boosting Tree (BT), SVR, and MLP to evaluate their effectiveness in predicting sand production from gas-hydrate-bearing sands. According to their empirical study, BT outperformed traditional methods in predictive accuracy and highlighted model selection and tuning as critical factors [120].

Fertig and colleagues (2022) studied quality prediction in milling processes with internal machine tool data and evaluated nine ML algorithms along with four DL architectures, finding ensemble methods (RF, Extra Trees) and deep networks (InceptionTime, ResNet) to be the most efficient for multivariate time-series classification. The researchers achieved precise quality monitoring by segmenting sensor data based on geometric characteristics, which led to high prediction accuracy and paved the way for machining process optimization [121]. Rubio-Loyola and colleagues (2022) targeted predictive modeling of black carbon emissions from industrial furnaces and trained their models with a dataset of 400,000 instances using logistic regression and Adaboost. Both predictive models performed exceptionally well, but logistic regression received preference for its reduced computational requirements. The research demonstrated the value of ML methods in monitoring emissions while providing operational insights and established the initial use of these techniques for predicting furnace emissions [122]. The research by Słowik and Urban (2022) demonstrated that their LSTM model provided better accuracy and computational efficiency for manufacturing microgrids' short-term energy consumption forecasting compared to both double-layer LSTM and CNN models while achieving a mean absolute error of 0.0464. The research confirmed that the model functioned effectively for immediate energy management tasks and grid operations [123].

The research by Nguyen et al. (2022) automated shot peening in remanufacturing through a distributed model predictive control system that used ML algorithms, including eXtreme Gradient Boosting (XGBoost) and sparse identification of nonlinear dynamics (SINDy). The proxy model successfully converted desired surface intensity into operational parameters, which resulted in improved accuracy and process stability while minimizing material waste [124]. Ibrahim et al. (2022) used ML and DL models such as XGBoost, ANN, and recurrent neural network (RNN) to accurately predict long-term oil, gas, and water production with R^2 values reaching 0.98 and demonstrated faster performance than traditional reservoir simulation methods. The models demonstrated high performance in oil and gas prediction but showed reduced accuracy in water production because of inconsistent data [125]. Li et al. (2022) implemented a probabilistic neural network with an enhanced particle swarm optimization algorithm to establish cost control classifications in manufacturing enterprises and assist with multi-objective decision-making. The study demonstrated practical success with their model achieving 96.1% classification accuracy and proving ML's utility for sustainability and stakeholder collaboration efforts [126].

In 2022, Stock et al. used ML techniques to enhance lithium-ion battery production through early detection and classification of quality attributes. The research utilized electrochemical impedance spectroscopy measurements along with cycling data from 29 pouch cells to develop linear regression models and ANNs. The ANN that demonstrated the highest performance reached a classification accuracy of 97% alongside a test error of 10.1%, which shows its promising capability for early quality prediction with limited observation time [127]. Park and Jeong (2022) developed a machine vision inspection system for mask production that used DL to enable defect detection and automate processes. The LAON PEOPLE NAVI AI Toolkit enabled their system to achieve a 61.1% productivity boost while elevating mask production from 2,600 to 4,189 masks per hour and lowering defect rates from 8.3% to 5.69%. The integration strengthened quality assurance while reducing workforce reliance and operational weariness [128]. The research team led by Filho introduced a predictive maintenance framework in 2022 that targets turbine maintenance in hydroelectric power facilities. The team processed 16 months of turbine load cycle data through RFs and ANNs models to attain prediction accuracies near 98% for 12- to 48-hour pre-failure windows. The researchers demonstrated how Most Influential Variables (MIV) could streamline models while maintaining original performance levels [129].

A time-domain process model developed by Ko (2022) simulated 2400 machining conditions and employed SVM with gradient tree boosting algorithms to classify end-milling stability with about 90% accuracy while proving that simulations could take the place of expensive experiments [130]. Gope et al. (2022) reviewed 440 historical datasets from polypropylene fiber melt spinning operations by applying DL and RF methods to enhance fiber quality and identify abnormal parameters, which resulted in the RF model correctly identifying quality-affecting settings with 100% accuracy [131]. Rodríguez et al. (2022) proposed a multi-agent deep RL method to carry out predictive maintenance on parallel machines, which resulted in a 60% increase in operational uptime and surpassed traditional maintenance strategies by 75% in efficiency [132].

Párizs et al. (2022) examined injection molding quality prediction using four ML classifiers—kNN, naïve Bayes, linear discriminant analysis (LDA), and decision tree—on pressure-based sensor data where decision trees reached the highest prediction accuracy (>90%) while requiring minimal computation time [133]. A study by Kannan et al. (2022) used ML to predict production delays in a Malaysian quarry company through five models that included neural networks and logistic regression. The research showed that neural networks and logistic regression models delivered peak performance with 96.8% accuracy and 97.3% F-measure, respectively, demonstrating how unplanned maintenance and over-utilization of specific machines affected these results [134]. Zhou et al. (2022) analyzed quality monitoring in resistance spot welding by integrating domain knowledge into ML pipelines and assessing the performance of linear regression, multi-layer perceptron, and SVR using real-world datasets from Bosch production lines. The research findings confirmed that prediction performance improved significantly through domain-informed feature engineering, which allowed even basic algorithms to work effectively when process-aware features provided guidance [135].

Mahdi et al. (2023) introduced a ML model for artificial lift selection in oil production that reached 93% prediction accuracy by examining six primary parameters across 24 oil wells and thus improved both selection efficiency and production output significantly [136]. Choi et al. (2023) constructed a Tap Temperature Prediction Model (TTPM) that uses SVR and analyzes data from 4598 EAFs heats to optimize energy use in the steel industry. Their model implementation achieved a 17% decrease in temperature deviation and saved 282 kWh of power per heat, which led to both economic benefits and reduced carbon emissions [137]. In 2023, Wang et al. introduced the Quality Filter (QF) model, which aims to stabilize product quality in the steel manufacturing sector. By combining ML and expert knowledge, their model optimized processing parameters through analysis of 128 samples of wind power steel and achieved an 82.81% classification accuracy while reducing product property fluctuations from raw material inconsistencies [138].

Santander et al. (2023) developed a detailed ML framework for refinery production planning that combines supervised and unsupervised learning techniques to handle model uncertainty as well as process disturbances and time correlations, which results in better performance compared to traditional methods [139]. McLaughlin and Choi (2023) developed five ML models using historical audit data to enhance accuracy in industrial compressed air system energy savings estimates and discovered that the distributed RF model achieved the highest prediction accuracy and improved energy efficiency by reducing auditor calculation errors [140]. The Laser Welder Predictive Maintenance Model (LW-PMM) created by Choi et al. (2023) combines autoencoders with LSTM for welder control data analysis to achieve high predictive maintenance accuracy that detects failures 27 hours ahead and saves facilities $23.5K yearly [141].

A multi-stage methodology proposed by Mehdiyev and his team in 2023 integrates Quantile Regression Forests to measure uncertainty along with SHapley Additive Explanations (SHAP) to enhance interpretability in predictive process monitoring. The model demonstrated high Prediction Interval Coverage Probability (~90%) and received validation through application in actual production planning scenarios, which proved its capability to enhance decision-making under uncertain conditions [142]. Abdelkader et al. (2023) investigated robustness challenges that arise during the implementation of ML models in production environments and underscored the importance of transparency and safety through standardized terminologies. They deployed protective patterns designed for ML and software systems, which increase model stability and trustworthiness in critical industrial applications [143]. Prabha et al. (2023) studied real-time data-driven modeling techniques that combined linear regression with decision trees to create fault forecasts while optimizing production parameters, including temperature and load. The study concluded that decision trees performed better than regression models, which supported improvements in predictive maintenance and operational efficiency [144].

Liu et al. (2023) targeted improvements in oil and gas production forecasts by utilizing static and dynamic ML models and highlighted data preprocessing and hyperparameter optimization as essential steps for better model performance. The research demonstrated that SVM and neural networks achieved better predictive accuracy and faster modeling performance than traditional methods [145]. Min et al. (2023) explored ML techniques for production prediction within the Coal Bed Methane (CBM) industry while concentrating on model interpretability alongside causal discovery. The combination of causal graphs with latent relationship analysis among geological, engineering, and treatment variables led to improved prediction accuracy between 5–31% while maintaining alignment with known CBM production mechanisms [146]. Kabir et al. (2023) optimized green hydrogen production processes through ML algorithms specifically targeting proton exchange membrane (PEM) and dark fermentation (DF) technologies for circular economy applications. The researchers used kNN and RF models to determine production efficiencies and identify crucial process parameters, including temperature and chemical oxygen demand, through permutation and partial dependency analyses [147].

Beni et al. (2023) applied ANN and Adaptive Neuro-Fuzzy Inference System (ANFIS) models to predict the energy efficiency and environmental footprint of almond and walnut farming processes and achieved outstanding R^2 values up to 0.997 using ANFIS, which displayed great potential for agricultural sustainability enhancement [148]. Du et al. (2023) developed a forecasting model for coalbed methane production using a bidirectional LSTM (Bi-LSTM) framework augmented with transfer learning techniques. The model achieved more than a 15% increase in prediction accuracy over conventional methods while keeping error margins below 200 m^3, which demonstrated its strength in small-sample applications and cross-well knowledge transfer abilities [149]. Singh et al. (2023) investigated biomass production optimization utilizing a decision tree algorithm to pinpoint key cultivation parameters for microalgae grown in wastewater. Experimental validation confirmed that the model achieved 81.25% accuracy with a 10% error margin, which proved both its predictive capability and scalability [150].

Tahir et al. (2023) applied ANNs to optimize hydrogen production from biomass gasification integrated with chemical looping by predicting syngas composition and determining optimal operating conditions, which resulted in a high R^2 value of 0.99 [151]. Yadav et al. (2023) created an optimized Synechococcus elongatus strain for increased isoprene production through an ANN coupled with a genetic algorithm (ANN-GA) system that optimized growth conditions and achieved a 29.52-fold production increase, which surpassed traditional statistical methods [152]. Cheng et al. (2023) used SVM, among other ML models, to predict green hydrogen production from photovoltaic (PV)-powered water electrolysis across different Chinese regions, which resulted in R^2 values above 0.95 and exhibited spatial variability in hydrogen potential performance superior to FbProphet [153].

In 2024, Zhao explored ML's transformative potential for predictive maintenance and design optimization while demonstrating how production efficiency gains through real-time data analysis integrated with IoT and cloud tech enable intelligent automation in mechanical systems [13]. Bost (2024) presented a structured framework that integrates ML into production engineering operations through a model register approach combined with data review and an ML opportunity matrix, which identifies suitable areas for ML application, including supervised learning to improve production optimization and turndown management as well as maintenance planning [2]. Wang et al. (2024) tackled production scheduling challenges in environments with high variability by creating the EMA-DCPM (dynamic critical path method) algorithm that used attention mechanisms and ML to accurately predict job times and enhance traditional critical path methods [154].

Naveed et al. (2024) explored responsible ML deployment challenges in dynamic production settings through their model-driven engineering technique, which monitors fairness, privacy, and safety during runtime. Through developing a meta-model and prototype, their research showed initial success in detecting human-centric requirement breaches but also recognized further development was necessary to adapt to changing operational requirements [155]. Bappy and Ahmed (2024) conducted empirical evaluations to assess the effects of ML on data collection methods applied to manufacturing and mechanical engineering fields. Their research demonstrated through 20 case studies and 15 industry reports that ML models such as neural networks and SVM achieved superior performance compared to traditional methods by reaching up to 95% accuracy improvement and reducing data collection time by 40% while decreasing defect detection errors and operational costs by 25% [14]. The research by Husom et al. (2024) expanded production engineering ML applications through the inclusion of environmental sustainability practices in the ML development lifecycle. The research team presented the CEMAI pipeline framework that processed sensor data from CNC machining and broaching operations across three industrial case studies to monitor ML development carbon emissions while optimizing their impact. Husom et al. (2024) showed how carbon-aware configurations can maintain model performance while lowering environmental impact [156].

Lemke et al. (2024) used ML models, including Ridge Regression, RF, and XGBoost, to predict egg production in poultry farming from multivariate time series data and found Ridge Regression achieved the lowest Mean Squared Error (MSE) of 19.74 and MAPE of 3.81% with a 7-day sliding window technique. The results showed how the model improved accuracy for agricultural production planning [157]. Yuan et al. (2024) researched ML-based modeling techniques to enhance hydrothermal and pyrolysis carbonization for biochar production from aquatic biomass materials. The researchers created a dataset of 586 data points to train tree-based models, where RF Regression obtained the highest accuracy for hydrochar-related outputs at R^2 values up to 0.98, while XGBoost performed well for pyrochar metrics. The research demonstrated both the importance of feedstock elemental composition for yield prediction and the effectiveness of their models by showing their ability to generalize through iterative learning [158]. Vats and Kumar (2024) used ML algorithms to study and improve nanoemulsion production through continuous ultrasonication. Their models achieved precise droplet size predictions by modifying the oil-to-surfactant ratio (OSR) and energy density and determined that an OSR of 4 maintained nanoemulsion stability for 50 days while supporting scalable production [159].

The team of D'Oelsnitz et al. (2024) engineered an Amaryllidaceae enzyme in Escherichia coli through the use of a structure-based ResNet named MutComputeX, which was fused with a biosensor system to improve biocatalytic efficiency in therapeutic alkaloid synthesis. The application of their strategy resulted in a 60% product titer rise together with a doubling of catalytic activity and a three-fold decrease in off-product formation, which demonstrated ML capabilities in molecular production enhancement [160]. Vuković et al. (2024) developed a multi-objective optimization system using local ML models for the ironmaking industry to investigate and control system interdependencies. A sintering plant case study revealed that process output improvements and operational cost reductions were achieved by quantifying system interdependencies through causality-informed ML models while demonstrating the importance of transparency and domain knowledge in industrial process optimization [161]. Cruz et al. (2024) introduced a fully automated ML methodology designed specifically for small and medium-sized enterprises (SMEs) that combines automated ML pipelines with multi-objective optimization features to create production models. The real manufacturing process saw a 3.19% productivity gain and a 2.15% defect rate drop when the researchers applied their methodology, which surpassed traditional trial-and-error techniques [162].

The research by Ross-Veitía et al. (2024) compared multiple regression-based ML models to forecast industrial steam boiler emissions (CO, CO_2, NO_x) from operational data and found Gradient Boosting Regression (GBR) to deliver superior results with a R^2 of 99.80% and a mean absolute error of 0.51, thus boosting combustion efficiency and environmental compliance [163]. Xu et al. (2024) developed a ML approach that combined molecular reaction mechanisms with big data processing to optimize chemical processes through multi-objective optimization, achieving simultaneous improvements in conversion rates (4.44%), product yields (4.25%), energy use (−6.23%), and greenhouse gas (GHG) emissions (−12.60%) [164]. Wang et al. (2024) used ANNs to develop a 6-bed, 12-step vacuum pressure swing adsorption (VPSA) system for coke oven gas hydrogen recovery, which achieved 99.99% hydrogen purity and up to 66.54% exergy efficiency while showing purity-recovery tradeoffs through Non-dominated Sorting Genetic Algorithm II (NSGA-II) optimization [165].

In 2024, Chu's research combined GPR with neural networks to model Z-axis displacement and detect sensor anomalies in machine tools for enhanced machining precision and operational stability [166]. Asiedu et al. (2024) conducted research on solar PV production forecasting using ML, which revealed that different models performed best at different forecast horizons: ANNs worked best for day-ahead predictions, while hybrid XGBoost–RF models excelled in week-ahead forecasts, and RFs performed best for long-term forecasts, showing the necessity of choosing models based on specific temporal forecasting needs [167]. Lahafdoozian et al. (2024) investigated hydrogen production from plastic waste through gasification processes by combining process simulations with ML techniques to optimize the process. The study demonstrated that RF models effectively predicted hydrogen yield with R^2 values above 0.99 and highlighted how temperature was the main factor for optimal cold gas efficiency and hydrogen production [168].

The study by Shahzad et al. (2024) applied both ANN and SVM to model solar-to-formic acid conversion and found SVM to exhibit superior predictive accuracy, which resulted in producing 2.25×10^6 g of formic acid from 2355 kW generated by a dual-axis PV tracker in Vietnam [169]. The Spatial Correlation-based ML Prediction Model (SC-MLPM) framework created by Yi et al. (2024) predicts shale gas locations in China's Sichuan Basin through an enhanced kNN and Huber loss-based XGBoost (Hu-XGB) model, which achieved a mean relative error of only 9.8% while accurately locating productive sweet spots [170]. Yang et al. (2024) investigated hydrogen production capabilities in 100 MW PV systems under Kucha conditions with seven ML models and found LSTM to be the top performer ($R^2 = 0.8402$), while also discovering heterojunction with intrinsic thin layer (HJT) PV technology as the most cost-effective method to reduce hydrogen production expenses [171]. The research team led by Elaziz et al. (2024) created the LSTM-election-based optimization algorithm (EBOA) model, which successfully predicted freshwater production in membrane desalination systems with near-perfect accuracy ($R^2 = 0.998$ training; 0.988 testing) while outperforming other optimized LSTM models in RMSE evaluations [172]. Leng et al. (2024) utilized GBR and RF algorithms to enhance hydrochar production from biomass hydrothermal carbonization and found GBR achieved better prediction accuracy with an average R^2 of 0.93 while identifying important operational parameters affecting yield and composition [173].

CHALLENGES AND FUTURE DIRECTIONS IN ADOPTING ML IN PRODUCTION ENGINEERING

Challenges

Organizations struggle to maximize the potential of ML solutions in production settings even though ML offers significant benefits.

- *Data Quality and Representativeness*: ML models show compromised reliability when they use inadequate training data with poor quality, unrepresentative samples, and imbalanced data sets [174]. Ongoing data validation combined with quality assurance processes remains an essential task [174].
- *Model Interpretability*: The opaque nature of ML models through DL approaches creates significant challenges for explainability and compliance in vital industrial applications [175].
- *Integration Complexities*: The integration of ML with existing industrial systems demands extensive technical knowledge and generates substantial expenses for compatibility with older infrastructure systems [64].
- *Operational Robustness*: ML deployments need to resist diverse data patterns along with changing process parameters and environmental conditions through strong and flexible model management techniques [143].
- *Workforce Skills and Cultural Change*: Personnel training along with cross-functional collaboration (MLOps) and development of a data-centric operational culture drive successful ML implementation in manufacturing [176].
- *Ethical, Regulatory, and Security Concerns*: The growing presence of AI and ML systems requires a proactive approach to security, transparency, and workforce challenges [68].

Future Directions

Progress in ML combined with fast digital transformation and Industrial IoT development defines the emerging production engineering paradigm.

- *Edge and Federated Learning*: Lightweight ML models are implemented on edge devices positioned on the production floor to support real-time decision-making by reducing latency [3].
- *Hybrid and Physics-Informed Models*: The integration of ML with conventional engineering simulations produces better accuracy and interpretability in modeling complex processes [19].
- *Autonomous and Collaborative Robotics*: Robots will utilize RL together with natural language processing to achieve greater autonomy and intelligent adaptation while ensuring safer human-robot collaboration [56].
- *Sustainability-Driven Analytics*: Manufacturing stages require advanced ML systems that focus on waste reduction and resource optimization to reach sustainability goals [57].
- *Human-in-the-Loop and Explainable AI*: Model transparency combined with workforce collaboration produces reliable and practical outcomes in essential industrial applications [175].

CONCLUSION

The study investigates how ML methods have been incorporated into production engineering processes. ML applications have led to major advancements in production processes through defect detection capabilities and optimization of manufacturing parameters. The application of ML results in improved product quality and greater process efficiency, according to multiple studies presented in the paper. The study demonstrates how predictive maintenance models have replaced traditional maintenance systems through the implementation of LSTM autoencoders. Real-time monitoring and machine data classification during this transition become essential for reducing downtime and boosting operational reliability. The review provides examples of multiple ML applications in production engineering that span predictive maintenance and quality control to resource allocation. Machine vision models demonstrate effectiveness in quality control by detecting defects and generating useful insights within Industry 4.0 applications. The significance of strong model validation and testing processes becomes evident as they ensure ML models perform well on new data and satisfy actual production requirements. Building stakeholder trust requires the adoption of comprehensive testing strategies along with prioritizing reproducibility and interpretability. Future research directions as identified by the paper include both algorithm refinement in ML as well as the discovery of new production engineering applications. The process requires selecting models that fit particular production situations while implementing explainable AI to improve transparency in decision-making. The detailed assessment reveals how ML can revolutionize production engineering and calls for ongoing research and application of these technologies to advance industrial efficiency and innovation.

REFERENCES

[1] Suthahar, P., Palanikumar, K., Ponshanmugakumar, A., & Anbuchezhiyan, G. (2024). Machine Learning Advancements in Machining Processes: A Comprehensive Review for Manufacturing Optimization. *Journal of Physics: Conference Series*, *2837*, 012102. https://doi.org/10.1088/1742-6596/2837/1/012102

[2] Bost, S. (2024). Machine Learning for Enhanced Production Optimisation and Management. *SPE Asia Pacific Oil and Gas Conference and Exhibition*, D021S012R002. https://doi.org/10.2118/221205-ms

[3] Manta-Costa, A., Araújo, S. O., Peres, R. S., & Barata, J. (2024). Machine learning applications in Manufacturing - challenges, trends, and future directions. *IEEE Open Journal of the Industrial Electronics Society*, *5*, 1085–1103. https://doi.org/10.1109/ojies.2024.3431240

[4] Chopra, V., & Priyadarshi, D. (2019). Role of machine learning in manufacturing sector. *International Journal of Recent Technology and Engineering (IJRTE)*, *8*(4), 2320–2328. https://doi.org/10.35940/ijrte.d8191.118419

[5] Ganjare, S. A., Satao, S. M., & Narwane, V. (2024). Systematic literature review of machine learning for manufacturing supply chain. *The TQM Journal*, *36*(8), 2236–2259. https://doi.org/10.1108/tqm-12-2022-0365

[6] Dyck, J. (2018). Machine learning for engineering. *2018 23rd Asia and South Pacific Design Automation Conference (ASP-DAC)*, 422–427. https://doi.org/10.1109/aspdac.2018.8297360

[7] Mutsuddi, S. (2023). Machine learning for predictive maintenance in manufacturing industries. *International Journal for Research in Applied Science and Engineering Technology*, *11*(4), 1175–1181. https://doi.org/10.22214/ijraset.2023.50098

[8] Shinkevich, A. I., & Malysheva, T. V. (2023). Diagnosis of core production dependence on production infrastructure based on ensemble method of machine learning. *Society: Politics, Economics, Law*, *8*, 114–120. https://doi.org/10.24158/pep.2023.8.14

[9] Kolhe, K., Somatkar, A. A., Bhandarkar, M. S., Kotangale, K. B., Ayane, S. S., & Shirke, S. I. (2023). Applications and challenges of machine learning techniques for smart manufacturing in Industry 4.0. *2023 7th International Conference on Computing, Communication, Control and Automation (ICCUBEA)*, 1–6. https://doi.org/10.1109/iccubea58933.2023.10392071

[10] Choudhury, M. R., & Dutta, A. (2022). Establishing Pertinence between Sorting Algorithms Prevailing in n log(n) time. *Journal of Robotics and Automation Research*, *3*(2), 220–226. https://doi.org/10.33140/jrar.03.02.09

[11] Maryam, A. H., Gehad, M. H., Abdelazziz, A. A., & Menatallah, A. H. (2024). Maximizing Value through Early Production Projects in Petroleum Engineering: Benefits, Challenges, and Key Success Factors. *ADIPEC*. https://doi.org/10.2118/222754-ms

[12] Wong, F. K., Sidek, S., Lee, S. S. F., Akkurt, R., Shi, Y., Tangtuengtin, P., Kalidas, S., Hazali, N. A., Mokhlis, M. M., Zulkipli, S. N. F., & Basarudin, M. A. (2024). Automated System for feature petrophysical Data Engineering – Improves ML behind casing opportunity results in highly deviated wells. In *Offshore Technology Conference Asia*. https://doi.org/10.4043/34871-ms

[13] Zhao, H. (2024). Machine learning analysis in mechanical engineering. *Computers and Artificial Intelligence.*, *1*(2), 1–5. https://doi.org/10.70267/2xfync82

[14] Bappy, M. A., & Ahmed, M. (2024). UTILIZING MACHINE LEARNING TO ASSESS DATA COLLECTION METHODS IN MANUFACTURING AND MECHANICAL ENGINEERING. *Academic Journal on Science, Technology, Engineering & Mathematics Education.*, *4*(2), 14–25. https://doi.org/10.69593/ajsteme.v4i02.73

[15] Biffl, S., Lüder, A., Rinker, F., Waltersdorfer, L., & Winkler, D. (2018). *Introducing engineering data logistics for production systems engineering*. Technical report CDL-SQI-2018-10, TU Wien. https://qse.ifs.tuwien.ac.at/wp-content/uploads/CDL-SQI-2018-10.pdf

[16] Al-Reyashi, M. H. A. (2023). EXPLORING THE EFFECTIVENESS OF DIFFERENT DATA CLEANING TECHNIQUES FOR IMPROVING DATA QUALITY IN MACHINE LEARNING. *Humanities & Natural Sciences Journal*, *4*(7), 149–153. https://doi.org/10.53796/hnsj4711

[17] Devarasetty, N. (2024). Optimizing data engineering for AI: improving data quality and preparation for machine learning application. *Research and Analysis Journal*, *7*(03), 1–29. https://doi.org/10.18535/raj.v7i03.397

[18] Deng, J. (2024). The data preprocessing technique in machine learning. *Science and Technology of Engineering, Chemistry and Environmental Protection*, *1*(10), 1–5. https://doi.org/10.61173/0hfqqz07

[19] Chakraborty, I., Kluk, D., & McNeill, S. (2020). Machine learning models with engineering insight: case studies from the oil and gas industry. *Proceedings of the ASME 2020 39th International Conference on Ocean, Offshore and Arctic Engineering. Volume 11: Petroleum Technology*, V011T11A003. https://doi.org/10.1115/omae2020-18895

[20] Khandelwal, A., & Awasthi, A. (2023). Data Engineering: The foundation for effective ML models. *Integrated Journal for Research in Arts and Humanities*, *3*(4), 235–247. https://doi.org/10.55544/ijrah.3.4.29

[21] Deepa, B., & Ramesh, K. (2021). Production level data pipeline environment for machine learning models. *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 404–407. https://doi.org/10.1109/icaccs51430.2021.9442035

[22] Aditya, N., Munthe, I. R., & Sihombing, V. (2024). A Comparative Analysis of Machine learning Algorithms for predicting paddy Production. *Sinkron : Jurnal Dan Penelitian Teknik Informatika*, *8*(2), 1200–1207. https://doi.org/10.33395/sinkron.v8i2.13666

[23] Chinecherem, O. O., Esther, U. N., & Chigbundu, K. (2023). Comprehensive review of supervised machine learning algorithms to identify the best and error free. *International Journal of Scholarly Research in Engineering and Technology*, *2*(1), 013–019. https://doi.org/10.56781/ijsret.2023.2.1.0028

[24] Alfarisi, O., Aung, Z., & Sassi, M. (2021). Deducing of Optimal Machine Learning Algorithms for Heterogeneity. *TechRxiv*. https://doi.org/10.36227/techrxiv.17162147.v1

[25] Lianzhai, D., Roestam, R., Sen, T. W., Fahmi, H., ChungKiat, O., & Hariyanto, D. T. (2024). Machine learning Algorithms for prediction of boiler steam production. *International Journal of Advances in Data and Information Systems*, *5*(2), 157–172. https://doi.org/10.59395/ijadis.v5i2.1339

[26] Sihombing, D. J. C. (2024). Application of feature engineering techniques and machine learning algorithms for property price prediction. *JITSI : Jurnal Ilmiah Teknologi Sistem Informasi*, *5*(2), 72–76. https://doi.org/10.62527/jitsi.5.2.241

[27] Katam, B. R. (2025). Advanced data engineering and machine learning approaches for accurate yield prediction in semiconductor manufacturing processes. *INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT*, *09*(01), 1–9. https://doi.org/10.55041/ijsrem40765

[28] Colyer, A. (2019). Putting Machine Learning into Production Systems: Data validation and software engineering for machine learning. *Queue*, *17*(4), 17–18. https://doi.org/10.1145/3358955.3365847

[29] Durelli, V. H. S., Monteiro, R., Durelli, R. S., Endo, A. T., Ferrari, F. C., & Souza, S. R. S. (2024). Property-based Testing for Machine Learning Models. *Anais Do IX Simpósio Brasileiro De Testes De Software Sistemático E Automatizado*, 39–48. https://doi.org/10.5753/sast.2024.3791

[30] Brower-Sinning, R., Lewis, G. A., Echeverría, S., & Ozkaya, I. (2024). Using Quality Attribute Scenarios for ML Model Test Case Generation. *2024 IEEE 21st International Conference on Software Architecture Companion (ICSA-C)*, 307–310. https://doi.org/10.1109/icsa-c63560.2024.00058

[31] Schwerdtner, P., Greßner, F., Kapoor, N., Assion, F., Sass, R., Günther, W., Hüger, F., & Schlicht, P. (2020). Risk assessment for machine learning models. *arXiv (Cornell University)*, 1–8. https://doi.org/10.48550/arxiv.2011.04328

[32] Murphy, B., Wakefield, A., & Friedman, J. (2008). Best Practices for Verification, Validation, and Test in Model-Based Design. *SAE Technical Papers on CD-ROM/SAE Technical Paper Series*. https://doi.org/10.4271/2008-01-1469

[33] Ghanta, S., Subramanian, S., Sundararaman, S., Khermosh, L., Sridhar, V., Arteaga, D., Luo, Q., Das, D., & Talagala, N. (2018). Interpretability and reproducability in production machine learning applications. *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 658–664. https://doi.org/10.1109/icmla.2018.00105

[34] Ananthi, S., B, L., M, S. C., & S, S. (2024). Framework for Platform Independent Machine Learning (ML) Model Execution. *2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)*, 728–732. https://doi.org/10.1109/idciot59759.2024.10467931

[35] Hanchuk, D. O., & Semerikov, S. O. (2025). Implementing MLOps practices for effective machine learning model deployment: A meta synthesis. *CEUR Workshop Proceedings*, 329–337. https://ceur-ws.org/Vol-3918/paper404.pdf

[36] Dharavath, M. (2024). Optimizing Machine Learning Models: A Data Engineering perspective. *International Journal for Multidisciplinary Research*, *6*(6). https://doi.org/10.36948/ijfmr.2024.v06i06.32242

[37] Singh, S. P. (2024). A Comprehensive Framework for ML Model Validation: From Development to Production Monitoring in Search and Recommendation Systems. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, *10*(6), 2146–2157. https://doi.org/10.32628/cseit2410612404

[38] Ghanta, S., Subramanian, S., Khermosh, L., Sundararaman, S., Shah, H., Goldberg, Y., Roselli, D. S., & Talagala, N. (2019). ML Health: Fitness tracking for production models. *arXiv (Cornell University)*, 1–13. https://doi.org/10.48550/arxiv.1902.02808

[39] Balasubramanian, A. (2020). End-to-end model lifecycle management: An MLOPS framework for drift detection, root cause analysis, and continuous retraining. *International Journal of Multidisciplinary Research and Growth Evaluation*, *1*(1), 92–102. https://doi.org/10.54660/.ijmrge.2020.1.1-92-102

[40] Jang, M., Georgiadis, A., Zhao, Y., & Silavong, F. (2024). DriftWatch: A Tool that Automatically Detects Data Drift and Extracts Representative Examples Affected by Drift. *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 6: Industry Track)*, 335–346. https://doi.org/10.18653/v1/2024.naacl-industry.28

[41] Pulicharla, M. R. (2019). Detecting and addressing model drift: Automated monitoring and real-time retraining in ML pipelines. *World Journal of Advanced Research and Reviews*, *3*(2), 147–152. https://doi.org/10.30574/wjarr.2019.3.2.0189

[42] Baumann, N., Kusmenko, E., Ritz, J., Rumpe, B., & Weber, M. B. (2022). Dynamic data management for continuous retraining. *Proceedings of the 25th International Conference on Model Driven Engineering Languages and Systems: Companion Proceedings*, 359–366. https://doi.org/10.1145/3550356.3561568

[43] Kanda, H., Okamura, H., & Dohi, T. (2023). A note on optimal retraining strategy for ML systems. *2023 10th International Conference on Dependable Systems and Their Applications (DSA)*, 730–733. https://doi.org/10.1109/dsa59317.2023.00104

[44] Gudepu, V., Chirumamilla, B., Chintapalli, V. R., Castoldi, P., Valcarenghi, L., Tamma, B. R., & Kondepu, K. (2024). Generative-AI for AI/ML model adaptive retraining in beyond 5G networks. *arXiv (Cornell University)*, 1–13. https://doi.org/10.48550/arxiv.2408.14827

[45] Mahadevan, A., & Mathioudakis, M. (2023). Cost-Effective retraining of machine learning models. *arXiv (Cornell University)*, 1–27. https://doi.org/10.48550/arxiv.2310.04216

[46] Sack, M. (2022). A data-centric reinforcement learning approach for self-updating machine learning models. *Proceedings Volume 12119, Open Architecture/Open Business Model Net-Centric Systems and Defense Transformation 2022*, 121190H. https://doi.org/10.1117/12.2615873

[47] Friese, S., Tout, K., & Kirby, G. (2014). *Best practices in ensuring data quality in Quality Rating and Improvement Systems (QRIS)*. Mathematica Policy Research Reports. https://www.acf.hhs.gov/sites/default/files/documents/opre/data_quality_brief_508_compliant.pdf

[48] Anusha, P., Kumar, S. S., & Pathipati, C. (2024). A Review of Machine Learning Strategies for Enhancing Efficiency and Innovation in Real-World Engineering Applications. *International Journal of Scientific Research in Engineering and Management*, *08*(06), 1–6. https://doi.org/10.55041/ijsrem35355

[49] Alzaidi, E. R. (2024). Improving industrial quality control by machine learning techniques. *Journal La Multiapp*, *5*(5), 692–711. https://doi.org/10.37899/journallamultiapp.v5i4.1537

[50] Erokhin, V. V., & Koryakin, A. R. (2023). MACHINE LEARNING FOR DEFECT DETECTION IN PRODUCT MANUFACTURING. *SOFT MEASUREMENTS AND COMPUTING*, *6*(67), 33–46. https://doi.org/10.36871/2618-9976.2023.06.004

[51] Shaikh, M. B., Patil, P. J., Thokal, P. V., & Pardeshi, D. B. (2024). Implementing machine learning for predictive maintenance in industrial machinery. *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 1–6. https://doi.org/10.1109/icccnt61001.2024.10724004

[52] Rani, A. J. M., Prithivirajan, V., Shiney, S. A., Babaiah, Ch., Swamy, H., & G, S. (2024). Enhancing Predictive Maintenance in Industrial IoT through Machine Learning Models. *2024 4th Asian Conference on Innovation in Technology (ASIANCON)*, 1–6. https://doi.org/10.1109/asiancon62057.2024.10837788

[53] Tebrizcik, S., Ersöz, S., & Aktepe, A. (2024). Improvement of quality performance in mask production by feature selection and machine learning methods and an application. *Savunma Bilimleri Dergisi*, *20*(1), 167–190. https://doi.org/10.17134/khosbd.1298163

[54] Gowda, D., V., Irshad, M. A., Mary, P. A., Prasad, Y. D., Prasad, K. D. V., & Tarigonda, H. (2024). Machine Learning-driven Quality Control Strategies for Industrial IoT Applications. *2024 4th Asian Conference on Innovation in Technology (ASIANCON)*, 1–7. https://doi.org/10.1109/asiancon62057.2024.10837745

[55] Dhabliya, D., Anandhasilambarasan, D., Singh, A., Waiker, V., Paonipagar, P., & Thomas, S. (2024). Leveraging machine learning in Real-Time automation for predictive maintenance in industry 4.0. *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 1–5. https://doi.org/10.1109/icccnt61001.2024.10725980

[56] Paul, K., Nafees, L., Hazra, A., & Ghosh, S. (2025). AI-Driven Industrial Robotics: Revolutionizing Automation with Machine Learning and Intelligent Adaptation. *International Journal for Multidisciplinary Research*, *7*(1), 1–9. https://doi.org/10.36948/ijfmr.2025.v07i01.37805

[57] Altarazi, F. (2024). Optimizing Waste Reduction in Manufacturing Processes Utilizing IoT Data with Machine Learning Approach for Sustainable Production. *Scalable Computing: Practice and Experience*, *25*(5), 4192–4204. https://doi.org/10.12694/scpe.v25i5.3191

[58] Moumeni, L., & Saber, M. (2024). Global Supply Chain: Enhance production cost efficiency through machine learning. *Journal of Computer Science*, *20*(9), 955–963. https://doi.org/10.3844/jcssp.2024.955.963

[59] Alkubaisi, S., Fenjan, A., Al-Sharify, T. A., Sadiq, S., Furaijl, H. B., & Ahmed, S. R. (2024). Machine learning models for optimizing supply chain management. *2024 8th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, 1–9. https://doi.org/10.1109/ismsit63511.2024.10757215

[60] Almamlook, R., Alshammaria, N., & Albalawi, O. (2022). Machine Learning for Predictive Maintenance in the Manufacturing Sector. *12th Annual International Conference on Industrial Engineering and Operations Management*, 3661–3662. https://doi.org/10.46254/an12.20220665

[61] Magena, C. (2024). Machine learning models for predictive maintenance in industrial engineering. *International Journal of Computing and Engineering*, *6*(3), 1–14. https://doi.org/10.47941/ijce.2137

[62] Smaoui, S., & Baklouti, M. (2024). ML-based failure detection approach for predictive maintenance in an industry 4.0 oriented web manufacturing control application. *2024 IEEE 7th International Conference on*

Advanced Technologies, Signal and Image Processing (ATSIP), 426–431. https://doi.org/10.1109/atsip62566.2024.10639020

[63] Sadraoui, Y., Er-Ratby, M., Kadiri, M. S., & Kobi, A. (2024). Quality process control and Preventive maintenance optimization of a grinding unit: a case study. *2024 10th International Conference on Optimization and Applications (ICOA)*, 1–6. https://doi.org/10.1109/icoa62581.2024.10753891

[64] Casalí, M. A. M., Gil, F. F., Boza, A., & Nazarenko, A. (2023). An industry maturity model for implementing Machine Learning operations in manufacturing. *International Journal of Production Management and Engineering*, *11*(2), 179–186. https://doi.org/10.4995/ijpme.2023.19138

[65] Trankov, M., Hadzhikolev, E., & Hadzhikoleva, S. (2024). MACHINE LEARNING ALGORITHMS IN QUALITY CONTROL OF TEXTILE FIBER MANUFACTURING. *Journal of Theoretical and Applied Information Technology*, *102*(4). http://www.jatit.org/volumes/Vol102No4/31Vol102No4.pdf

[66] Kakolu, S., & Faheem, M. A. (2024). Advanced robotics automation for field operations: Leveraging AI and machine learning for dynamic, real-world applications. *International Journal of Science and Research Archive*, *11*(1), 2605–2614. https://doi.org/10.30574/ijsra.2024.11.1.0010

[67] Pelleti, S., Kumar, J. R. R., Krushnasamy, V. S., & Raj, G. B. M. (2024). Intelligent Control Systems for Industrial Automation and Robotics. *2024 Ninth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)*, 1–5. https://doi.org/10.1109/iconstem60960.2024.10568781

[68] Singh, M., & Khan, S. A. L. A. (2024). Advances in autonomous robotics: integrating AI and machine learning for enhanced automation and control in industrial applications. *International Journal for Multidimensional Research Perspectives*, *2*(4), 74–90. https://doi.org/10.61877/ijmrp.v2i4.135

[69] Mitra, M. (2019). Editorial on advances in machine learning and robotics. *American Research Journal of Electronics and Communication Engineering*, *1*(1), 16–20. https://doi.org/10.21694/2643-3486.19003

[70] Wyrembek, M., Baryannis, G., & Brintrup, A. (2025). Causal machine learning for supply chain risk prediction and intervention planning. *International Journal of Production Research*, 1–20. https://doi.org/10.1080/00207543.2025.2458121

[71] IBM Data and AI Team. (2023). MLOps and the evolution of data science. *IBM*. Retrieved December 23, 2024, from https://www.ibm.com/think/topics/mlops

[72] Kazmierczak, J., Salama, K., Huerta, V., & Bahadur, S. K. J. (2024). *MLOps: Continuous delivery and automation pipelines in machine learning*. Google Cloud. Retrieved December 23, 2024, from https://cloud.google.com/architecture/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning

[73] Amazon Web Services, Inc. (2022). *Amazon Robotics Uses Amazon SageMaker and AWS Inferentia to Enable ML Inferencing at Scale*. Aws. Retrieved December 23, 2024, from https://aws.amazon.com/solutions/case-studies/amazon-robotics-case-study/

[74] Amershi, S., Begel, A., Bird, C., DeLine, R., Gall, H., Kamar, E., Nagappan, N., Nushi, B., & Zimmermann, T. (2019a). Software Engineering for Machine Learning: A Case Study. *2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP)*, 291–300. https://doi.org/10.1109/icse-seip.2019.00042

[75] Amershi, S., Begel, A., Bird, C., DeLIne, R., Gall, H., Kamar, E., Nagappan, N., Nushi, B., & Zimmermann, T. (2019b). *Software Engineering for Machine Learning: A Case Study*. Microsoft Research. Retrieved

December 23, 2024, from https://www.microsoft.com/en-us/research/publication/software-engineering-for-machine-learning-a-case-study/

[76] Berg, D. J., Cledat, R., Seeley, K., Srikanth, S., Wang, C., & Yu, D. (2024). *Supporting Diverse ML Systems at Netflix*. Netflix Technology Blog. Retrieved December 23, 2024, from https://netflixtechblog.com/supporting-diverse-ml-systems-at-netflix-2d2e6b6d205d

[77] Soto, J. A. C., Tavakolizadeh, F., & Gyulai, D. (2019). An online machine learning framework for early detection of product failures in an Industry 4.0 context. *International Journal of Computer Integrated Manufacturing*, *32*(4–5), 452–465. https://doi.org/10.1080/0951192x.2019.1571238

[78] Paul, A., Mozaffar, M., Yang, Z., Liao, W.-K., Choudhary, A., Cao, J., & Agrawal, A. (2019). A Real-Time Iterative Machine Learning Approach for Temperature Profile Prediction in Additive Manufacturing Processes. *2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, 541–550. https://doi.org/10.1109/dsaa.2019.00069

[79] Rodríguez, G. G., Gonzalez-Cava, J. M., & Pérez, J. A. M. (2020). An intelligent decision support system for production planning based on machine learning. *Journal of Intelligent Manufacturing*, *31*, 1257–1273. https://doi.org/10.1007/s10845-019-01510-y

[80] Dučić, N., Jovičić, A., Manasijević, S., Radiša, R., Ćojbašić, Ž., & Savković, B. (2020). Application of Machine Learning in the Control of Metal Melting Production Process. *Applied Sciences*, *10*(17), 1–15. https://doi.org/10.3390/app10176048

[81] San-Payo, G., Ferreira, J. C., Santos, P., & Martins, A. L. (2020). Machine learning for quality control system. *Journal of Ambient Intelligence and Humanized Computing*, *11*, 4491–4500. https://doi.org/10.1007/s12652-019-01640-4

[82] Gaikwad, A., Giera, B., Guss, G. M., Forien, J.-B., Matthews, M. J., & Rao, P. (2020). Heterogeneous sensing and scientific machine learning for quality assurance in laser powder bed fusion – A single-track study. *Additive Manufacturing*, *36*, 101659. https://doi.org/10.1016/j.addma.2020.101659

[83] Kaya, A., Keçeli, A. S., Catal, C., & Tekinerdogan, B. (2020). Sensor failure tolerable Machine Learning-Based Food Quality Prediction Model. *Sensors*, *20*(11), 3173. https://doi.org/10.3390/s20113173

[84] Orrù, P. F., Zoccheddu, A., Sassu, L., Mattia, C., Cozza, R., & Arena, S. (2020). Machine learning approach using MLP and SVM algorithms for the fault prediction of a centrifugal pump in the oil and gas industry. *Sustainability*, *12*(11), 4776. https://doi.org/10.3390/su12114776

[85] Arellano-Espitia, F., Delgado-Prieto, M., Martinez-Viol, V., Saucedo-Dorantes, J. J., & Osornio-Rios, R. A. (2020). Deep-Learning-Based methodology for fault diagnosis in electromechanical systems. *Sensors*, *20*(14), 3949. https://doi.org/10.3390/s20143949

[86] Lutz, B., Kisskalt, D., Mayr, A., Regulin, D., Pantano, M., & Franke, J. (2021). In-situ identification of material batches using machine learning for machining operations. *Journal of Intelligent Manufacturing*, *32*, 1485–1495. https://doi.org/10.1007/s10845-020-01718-3

[87] Khayyati, S., & Tan, B. (2021). A machine learning approach for implementing data-driven production control policies. *International Journal of Production Research*, *60*(10), 3107–3128. https://doi.org/10.1080/00207543.2021.1910872

[88] Sauter, E., Sarikaya, E., Winter, M., & Wegener, K. (2021). In-process detection of grinding burn using machine learning. *The International Journal of Advanced Manufacturing Technology*, *115*, 2281–2297. https://doi.org/10.1007/s00170-021-06896-9

[89] Rai, R., Tiwari, M. K., Ivanov, D., & Dolgui, A. (2021). Machine learning in manufacturing and industry 4.0 applications. *International Journal of Production Research*, *59*(16), 4773–4778. https://doi.org/10.1080/00207543.2021.1956675

[90] Simeone, A., Zeng, Y., & Caggiano, A. (2021). Intelligent decision-making support system for manufacturing solution recommendation in a cloud framework. *The International Journal of Advanced Manufacturing Technology*, *112*, 1035–1050. https://doi.org/10.1007/s00170-020-06389-1

[91] Xin, D., Miao, H., Parameswaran, A., & Polyzotis, N. (2021). Production Machine Learning Pipelines: Empirical Analysis and Optimization Opportunities. *SIGMOD '21: Proceedings of the 2021 International Conference on Management of Data*, 2639–2652. https://doi.org/10.1145/3448016.3457566

[92] Kang, Z., Catal, C., & Tekinerdogan, B. (2021). Remaining Useful life (RUL) prediction of equipment in production lines using artificial neural networks. *Sensors*, *21*(3), 932. https://doi.org/10.3390/s21030932

[93] Burggraf, P., Wagner, J., Heinbach, B., & Steinberg, F. (2021). Machine Learning-Based Prediction of missing components for assembly – A case study at an Engineer-to-Order manufacturer. *IEEE Access*, *9*, 105926–105938. https://doi.org/10.1109/access.2021.3075620

[94] Bampoula, X., Siaterlis, G., Nikolakis, N., & Alexopoulos, K. (2021). A deep learning model for predictive maintenance in Cyber-Physical production systems using LSTM autoencoders. *Sensors*, *21*(3), 972. https://doi.org/10.3390/s21030972

[95] Benbarrad, T., Salhaoui, M., Kenitar, S. B., & Arioua, M. (2021). Intelligent Machine Vision model for defective product inspection based on machine learning. *Journal of Sensor and Actuator Networks*, *10*(1), 7. https://doi.org/10.3390/jsan10010007

[96] Siler, D. L., Pepin, J. D., Vesselinov, V. V., Mudunuru, M. K., & Ahmmed, B. (2021). Machine learning to identify geologic factors associated with production in geothermal fields: a case-study using 3D geologic data, Brady geothermal field, Nevada. *Geothermal Energy*, *9*, 17. https://doi.org/10.1186/s40517-021-00199-8

[97] Szajna, A., Kostrzewski, M., Ciebiera, K., Stryjski, R., & Woźniak, W. (2021). Application of the deep CNN-Based method in industrial system for wire marking identification. *Energies*, *14*(12), 3659. https://doi.org/10.3390/en14123659

[98] Wang, B., Sharma, J., Chen, J., & Persaud, P. (2021). Ensemble Machine Learning Assisted Reservoir Characterization Using Field Production Data–An Offshore Field Case Study. *Energies*, *14*(4), 1052. https://doi.org/10.3390/en14041052

[99] Wang, H., Qiao, L., Lu, S., Chen, F., Fang, Z., He, X., Zhang, J., & He, T. (2021). A novel shale gas production prediction model based on machine learning and its application in optimization of multistage fractured horizontal wells. *Frontiers in Earth Science*, *9*. https://doi.org/10.3389/feart.2021.726537

[100] Leon-Medina, J. X., Camacho, J., Gutierrez-Osorio, C., Salomón, J. E., Rueda, B., Vargas, W., Sofrony, J., Restrepo-Calle, F., Pedraza, C., & Tibaduiza, D. (2021). Temperature prediction using multivariate time series deep learning in the lining of an electric arc furnace for ferronickel production. *Sensors*, *21*(20), 6894. https://doi.org/10.3390/s21206894

[101] Yoo, S., & Kang, N. (2021). Explainable artificial intelligence for manufacturing cost estimation and machining feature visualization. *Expert Systems With Applications*, *183*, 115430. https://doi.org/10.1016/j.eswa.2021.115430

[102] Moretti, E., Tappia, E., Limère, V., & Melacini, M. (2021). Exploring the application of machine learning to the assembly line feeding problem. *Operations Management Research*, *14*, 403–419. https://doi.org/10.1007/s12063-021-00201-3

[103] Griffin, J. M., Shanbhag, V. V., Pereira, M. P., & Rolfe, B. F. (2021). Application of machine learning for acoustic emissions waveform to classify galling wear on sheet metal stamping tools. *The International Journal of Advanced Manufacturing Technology*, *116*, 579–596. https://doi.org/10.1007/s00170-021-07408-5

[104] Badarinath, P. V., Chierichetti, M., & Kakhki, F. D. (2021). A machine learning approach as a surrogate for a finite element analysis: status of research and application to one dimensional systems. *Sensors*, *21*(5), 1654. https://doi.org/10.3390/s21051654

[105] Stauder, M., & Kühl, N. (2022). AI for in-line vehicle sequence controlling: development and evaluation of an adaptive machine learning artifact to predict sequence deviations in a mixed-model production line. *Flexible Services and Manufacturing Journal*, *34*, 709–747. https://doi.org/10.1007/s10696-021-09430-x

[106] Garouani, M., Ahmad, A., Bouneffa, M., Hamlich, M., Bourguin, G., & Lewandowski, A. (2022). Using meta-learning for automated algorithms selection and configuration: an experimental framework for industrial big data. *Journal of Big Data*, *9*, 57. https://doi.org/10.1186/s40537-022-00612-4

[107] Tan, L. B., & Nhat, N. D. P. (2022). Prediction and optimization of process parameters for composite thermoforming using a machine learning approach. *Polymers*, *14*(14), 2838. https://doi.org/10.3390/polym14142838

[108] Zhang, B., Hua, J., Cai, L., Gao, Y., & Li, Y. (2022). Optimization of production parameters of particle gluing on internal bonding strength of particleboards using machine learning technology. *Journal of Wood Science*, *68*, 21. https://doi.org/10.1186/s10086-022-02029-2

[109] Grishina, Y., & Sorokin, D. (2022). Local-to-global learning for iterative training of production SLU models on new features. *ACL Anthology*, 103–111. https://doi.org/10.18653/v1/2022.naacl-industry.13

[110] Salem, A. M., Yakoot, M. S., & Mahmoud, O. (2022). A novel machine learning model for autonomous analysis and diagnosis of well integrity failures in artificial-lift production systems. *ADVANCES IN GEO-ENERGY RESEARCH*, *6*(2), 123–142. https://doi.org/10.46690/ager.2022.02.05

[111] Hojati, F., Azarhoushang, B., Daneshi, A., & Khiabani, R. H. (2022). Prediction of machining condition using time series imaging and deep learning in slot milling of titanium alloy. *Journal of Manufacturing and Materials Processing*, *6*(6), 145. https://doi.org/10.3390/jmmp6060145

[112] Natanael, D., & Sutanto, H. (2022). Machine Learning Application using Cost-Effective Components for Predictive Maintenance in Industry: A tube filling machine case study. *Journal of Manufacturing and Materials Processing*, *6*(5), 108. https://doi.org/10.3390/jmmp6050108

[113] Tama, B. A., Vania, M., Kim, I., & Lim, S. (2022). An EfficientNet-Based weighted ensemble model for industrial machine malfunction detection using acoustic signals. *IEEE Access*, *10*, 34625–34636. https://doi.org/10.1109/access.2022.3160179

[114] Cao, Q., Zanni-Merk, C., Samet, A., Reich, C., De Bertrand De Beuvron, F., Beckmann, A., & Giannetti, C. (2022). KSPMI: a knowledge-based system for predictive maintenance in industry 4.0. *Robotics and Computer-Integrated Manufacturing*, *74*, 102281. https://doi.org/10.1016/j.rcim.2021.102281

[115] Motz, M., Krauß, J., & Schmitt, R. H. (2022). Benchmarking of hyperparameter optimization techniques for machine learning applications in production. *Advances in Industrial and Manufacturing Engineering*, *5*, 100099. https://doi.org/10.1016/j.aime.2022.100099

[116] Feng, M., & Li, Y. (2022). Predictive maintenance decision making based on reinforcement learning in multistage production systems. *IEEE Access*, *10*, 18910–18921. https://doi.org/10.1109/access.2022.3151170

[117] Bhowmik, P. (2022). Machine Learning in Production: From Experimented ML Model to System. *Journal of Robotics and Automation Research*, *3*(2), 200–208. https://doi.org/10.33140/jrar.03.02.08

[118] Rom, M., Brockmann, M., Herty, M., & Iacomini, E. (2022). Machine learning tools in production engineering. *The International Journal of Advanced Manufacturing Technology*, *121*, 4793–4804. https://doi.org/10.1007/s00170-022-09591-5

[119] Bulaev, V. I. (2022). FEATURES OF INTRODUCING MACHINE LEARNING MODELS INTO PRODUCTION. *GEOPHYSICS*, *4*, 143–147. https://doi.org/10.34926/geo.2022.72.95.001

[120] Song, J., Li, Y., Liu, S., Xiong, Y., Pang, W., He, Y., & Mu, Y. (2022). Comparison of machine learning algorithms for sand production prediction: An example for a Gas-Hydrate-Bearing sand case. *Energies*, *15*(18), 6509. https://doi.org/10.3390/en15186509

[121] Fertig, A., Weigold, M., & Chen, Y. (2022). Machine Learning based quality prediction for milling processes using internal machine tool data. *Advances in Industrial and Manufacturing Engineering*, *4*, 100074. https://doi.org/10.1016/j.aime.2022.100074

[122] Rubio-Loyola, J., & Paul-Fils, W. R. S. (2022). Applied Machine Learning in Industry 4.0: Case-Study research in predictive models for black carbon emissions. *Sensors*, *22*(10), 3947. https://doi.org/10.3390/s22103947

[123] Slowik, M., & Urban, W. (2022). Machine Learning Short-Term energy consumption forecasting for microgrids in a manufacturing plant. *Energies*, *15*(9), 3382. https://doi.org/10.3390/en15093382

[124] Nguyen, V. B., Teo, A., Ba, T., Ahluwalia, K., & Kang, C. W. (2022). A distributed model predictive control with machine learning for automated shot peening machine in remanufacturing processes. *The International Journal of Advanced Manufacturing Technology*, *122*, 2419–2431. https://doi.org/10.1007/s00170-022-10018-4

[125] Ibrahim, N. M., Alharbi, A. A., Alzahrani, T. A., Abdulkarim, A. M., Alessa, I. A., Hameed, A. M., Albabtain, A. S., Alqahtani, D. A., Alsawwaf, M. K., & Almuqhim, A. A. (2022). Well performance Classification and prediction: Deep learning and machine learning long term regression experiments on oil, gas, and water production. *Sensors*, *22*(14), 5326. https://doi.org/10.3390/s22145326

[126] Li, X., Ma, C., & Lv, Y. (2022). Environmental Cost Control of Manufacturing Enterprises via Machine Learning under Data Warehouse. *Sustainability*, *14*(18), 11571. https://doi.org/10.3390/su141811571

[127] Stock, S., Pohlmann, S., Günter, F. J., Hille, L., Hagemeister, J., & Reinhart, G. (2022). Early quality classification and prediction of battery cycle life in production using machine learning. *Journal of Energy Storage*, *50*, 104144. https://doi.org/10.1016/j.est.2022.104144

[128] Park, M., & Jeong, J. (2022). Design and implementation of Machine Vision-Based Quality Inspection System in mask manufacturing process. *Sustainability*, *14*(10), 6009. https://doi.org/10.3390/su14106009

[129] De Aguiar Vallim Filho, A. R., Moraes, D. F., De Aguiar Vallim, M. V. B., Da Silva, L. S., & Da Silva, L. A. (2022). A Machine Learning modeling framework for predictive maintenance based on equipment load Cycle: an application in a real world case. *Energies*, *15*(10), 3724. https://doi.org/10.3390/en15103724

[130] Ko, J. H. (2022). Machining stability categorization and prediction using process model guided machine learning. *Metals*, *12*(2), 298. https://doi.org/10.3390/met12020298

[131] Gope, A. K., Liao, Y.-S., & Kuo, C.-F. J. (2022). Quality prediction and abnormal processing parameter identification in polypropylene fiber melt spinning using artificial intelligence machine learning and deep learning algorithms. *Polymers*, *14*(13), 2739. https://doi.org/10.3390/polym14132739

[132] Rodríguez, M. L. R., Kubler, S., De Giorgio, A., Cordy, M., Robert, J., & Traon, Y. L. (2022). Multi-agent deep reinforcement learning based Predictive Maintenance on parallel machines. *Robotics and Computer-Integrated Manufacturing*, *78*, 102406. https://doi.org/10.1016/j.rcim.2022.102406

[133] Párizs, R. D., Török, D., Ageyeva, T., & Kovács, J. G. (2022). Machine learning in Injection Molding: an industry 4.0 method of quality prediction. *Sensors*, *22*(7), 2704. https://doi.org/10.3390/s22072704

[134] Kannan, R., Halim, H. A. A., Ramakrishnan, K., Ismail, S., & Wijaya, D. R. (2022). Machine learning approach for predicting production delays: a quarry company case study. *Journal of Big Data*, *9*, 94. https://doi.org/10.1186/s40537-022-00644-w

[135] Zhou, B., Pychynski, T., Reischl, M., Kharlamov, E., & Mikut, R. (2022). Machine learning with domain knowledge for predictive quality monitoring in resistance spot welding. *Journal of Intelligent Manufacturing*, *33*, 1139–1163. https://doi.org/10.1007/s10845-021-01892-y

[136] Mahdi, M. A. A., Amish, M., & Oluyemi, G. (2023). An Artificial lift selection approach using Machine Learning: a case study in Sudan. *Energies*, *16*(6), 2853. https://doi.org/10.3390/en16062853

[137] Choi, S.-W., Seo, B.-G., & Lee, E.-B. (2023). Machine Learning-Based tap temperature prediction and control for optimized power consumption in stainless electric arc furnaces (EAF) of steel plants. *Sustainability*, *15*(8), 6393. https://doi.org/10.3390/su15086393

[138] Wang, X., Li, H., Pan, T., Su, H., & Meng, H. (2023). Material Quality Filter Model: Machine Learning Integrated with Expert Experience for Process Optimization. *Metals*, *13*(5), 898. https://doi.org/10.3390/met13050898

[139] Santander, O., Kuppuraj, V., Harrison, C. A., & Baldea, M. (2023). A machine learning framework for improving refinery production planning. *AIChE Journal*, *69*(8), e18109. https://doi.org/10.1002/aic.18109

[140] McLaughlin, E., & Choi, J.-K. (2023). Utilizing machine learning models to estimate energy savings from an industrial energy system. *Resources, Environment and Sustainability*, *12*, 100103. https://doi.org/10.1016/j.resenv.2022.100103

[141] Choi, J.-S., Choi, S.-W., & Lee, E.-B. (2023). Modeling of Predictive Maintenance Systems for Laser-Welders in Continuous Galvanizing Lines Based on Machine Learning with Welder Control Data. *Sustainability*, *15*(9), 7676. https://doi.org/10.3390/su15097676

[142] Mehdiyev, N., Majlatow, M., & Fettke, P. (2023). Quantifying and Explaining Machine Learning Uncertainty in Predictive Process Monitoring: An Operations Research Perspective. *arXiv (Cornell University)*, 1–40. https://doi.org/10.48550/arxiv.2304.06412

[143] Abdelkader, H., Abdelrazek, M., Schneider, J., Rani, P., & Vasa, R. (2023). Robustness Attributes to Safeguard Machine Learning Models in Production. *2023 IEEE Engineering Informatics*, 1–9. https://doi.org/10.1109/ieeeconf58110.2023.10520555

[144] Prabha, S. C., Sivaraaj, P., & Lakshmi, S. K. (2023). Data analysis and machine learning-based modeling for real-time production. *THE SCIENTIFIC TEMPER*, *14*(03), 637–640. https://doi.org/10.58414/scientifictemper.2023.14.3.11

[145] Liu, Z., Li, S., & Li, L. (2023). Research on oil and gas production prediction process based on machine learning. *International Journal of Energy*, *2*(2), 76–79. https://doi.org/10.54097/ije.v2i2.7773

[146] Min, C., Wen, G., Gou, L., Li, X., & Yang, Z. (2023). Interpretability and causal discovery of the machine learning models to predict the production of CBM wells after hydraulic fracturing. *Energy*, *285*, 129211. https://doi.org/10.1016/j.energy.2023.129211

[147] Kabir, M. M., Roy, S. K., Alam, F., Nam, S. Y., Im, K. S., Tijing, L., & Shon, H. K. (2023). Machine learning-based prediction and optimization of green hydrogen production technologies from water industries for a circular economy. *Desalination*, *567*, 116992. https://doi.org/10.1016/j.desal.2023.116992

[148] Beni, M. S., Parashkoohi, M. G., Beheshti, B., Ghahderijani, M., & Bakhoda, H. (2023). Application of machine learning to predict of energy use efficiency and damage assessment of almond and walnut production. *Environmental and Sustainability Indicators*, *20*, 100298. https://doi.org/10.1016/j.indic.2023.100298

[149] Du, S., Wang, M., Yang, J., Zhao, Y., Wang, J., Yue, M., Xie, C., & Song, H. (2023). An enhanced prediction framework for coalbed methane production incorporating deep learning and transfer learning. *Energy*, *282*, 128877. https://doi.org/10.1016/j.energy.2023.128877

[150] Singh, V., Verma, M., Chivate, M. S., & Mishra, V. (2023). Machine learning-based optimisation of microalgae biomass production by using wastewater. *Journal of Environmental Chemical Engineering*, *11*(6), 111387. https://doi.org/10.1016/j.jece.2023.111387

[151] Tahir, F., Arshad, M. Y., Saeed, M. A., & Ali, U. (2023). Integrated process for simulation of gasification and chemical looping hydrogen production using Artificial Neural Network and machine learning validation. *Energy Conversion and Management*, *296*, 117702. https://doi.org/10.1016/j.enconman.2023.117702

[152] Yadav, I., Rautela, A., Gangwar, A., Wagadre, L., Rawat, S., & Kumar, S. (2023). Enhancement of isoprene production in engineered Synechococcus elongatus UTEX 2973 by metabolic pathway inhibition and machine learning-based optimization strategy. *Bioresource Technology*, *387*, 129677. https://doi.org/10.1016/j.biortech.2023.129677

[153] Cheng, G., Luo, E., Zhao, Y., Yang, Y., Chen, B., Cai, Y., Wang, X., & Dong, C. (2023). Analysis and prediction of green hydrogen production potential by photovoltaic-powered water electrolysis using machine learning in China. *Energy*, *284*, 129302. https://doi.org/10.1016/j.energy.2023.129302

[154] Wang, L., Liu, H., Xia, M., Wang, Y., & Li, M. (2024). A machine learning based EMA-DCPM algorithm for production scheduling. *Scientific Reports*, *14*, 1–15. https://doi.org/10.1038/s41598-024-71355-w

[155] Naveed, H., Grundy, J., Arora, C., Khalajzadeh, H., & Haggag, O. (2024). Towards Runtime Monitoring for Responsible Machine Learning using Model-driven Engineering. *MODELS '24: Proceedings of the ACM/IEEE 27th International Conference on Model Driven Engineering Languages and Systems*, 195–202. https://doi.org/10.1145/3640310.3674092

[156] Husom, E. J., Sen, S., & Goknil, A. (2024). Engineering Carbon Emission-aware Machine Learning Pipelines. *CAIN 2024: Proceedings of the IEEE/ACM 3rd International Conference on AI Engineering - Software Engineering for AI*, *30*, 118–128. https://doi.org/10.1145/3644815.3644943

[157] Lemke, V., Andrade, J. O., & Komati, K. S. (2024). Machine Learning Techniques for Egg Production Prediction. In *Ibero-Latin American Congress on Computational Methods in Engineering (CILAMCE)*. https://doi.org/10.55592/cilamce.v6i06.10159

[158] Yuan, Z., Wang, Y., Zhu, L., Zhang, C., & Sun, Y. (2024). Machine-learning-aided biochar production from aquatic biomass. *Carbon Research*, *3*, 77. https://doi.org/10.1007/s44246-024-00169-2

[159] Vats, D., & Kumar, V. (2024). Enhanced Nanoemulsion Engineering through Continuous Ultrasonication and Integrated Machine Learning Algorithms. *Industrial & Engineering Chemistry Research*, *64*(1), 357–368. https://doi.org/10.1021/acs.iecr.4c02489

[160] D'Oelsnitz, S., Diaz, D. J., Kim, W., Acosta, D. J., Dangerfield, T. L., Schechter, M. W., Minus, M. B., Howard, J. R., Do, H., Loy, J. M., Alper, H. S., Zhang, Y. J., & Ellington, A. D. (2024). Biosensor and machine learning-aided engineering of an amaryllidaceae enzyme. *Nature Communications*, *15*, 2084. https://doi.org/10.1038/s41467-024-46356-y

[161] Vuković, M., Koutroulis, G., Mutlu, B., Krahwinkler, P., & Thalmann, S. (2024). Local machine learning model-based multi-objective optimization for managing system interdependencies in production: A case study from the ironmaking industry. *Engineering Applications of Artificial Intelligence*, *133*, 108099. https://doi.org/10.1016/j.engappai.2024.108099

[162] Cruz, Y. J., Villalonga, A., Castaño, F., Rivas, M., & Haber, R. E. (2024). Automated machine learning methodology for optimizing production processes in small and medium-sized enterprises. *Operations Research Perspectives*, *12*, 100308. https://doi.org/10.1016/j.orp.2024.100308

[163] Ross-Veitía, B. D., Palma-Ramírez, D., Arias-Gilart, R., Conde-García, R. E., Espinel-Hernández, A., Nuñez-Alvarez, J. R., Hernández-Herrera, H., & Llosas-Albuerne, Y. E. (2024). Machine learning regression algorithms to predict emissions from steam boilers. *Heliyon*, *10*(5), e26892. https://doi.org/10.1016/j.heliyon.2024.e26892

[164] Xu, W., Wang, Y., Zhang, D., Yang, Z., Yuan, Z., Lin, Y., Yan, H., Zhou, X., & Yang, C. (2024). Transparent AI-assisted chemical engineering process: Machine learning modeling and multi-objective optimization for integrating process data and molecular-level reaction mechanisms. *Journal of Cleaner Production*, *448*, 141412. https://doi.org/10.1016/j.jclepro.2024.141412

[165] Wang, J., Chen, X., Liu, L., Du, T., Webley, P. A., & Li, G. K. (2024). Vacuum pressure swing adsorption intensification by machine learning: Hydrogen production from coke oven gas. *International Journal of Hydrogen Energy*, *69*, 837–854. https://doi.org/10.1016/j.ijhydene.2024.05.100

[166] Chu, W.-L. (2024). Integrating machine learning and feature analysis for predicting and managing thermal deformation in machine tools. *Case Studies in Thermal Engineering*, *57*, 104343. https://doi.org/10.1016/j.csite.2024.104343

[167] Asiedu, S. T., Nyarko, F. K. A., Boahen, S., Effah, F. B., & Asaaga, B. A. (2024). Machine learning forecasting of solar PV production using single and hybrid models over different time horizons. *Heliyon*, *10*(7), e28898. https://doi.org/10.1016/j.heliyon.2024.e28898

[168] Lahafdoozian, M., Khoshkroudmansouri, H., Zein, S. H., & Jalil, A. A. (2024). Hydrogen production from plastic waste: A comprehensive simulation and machine learning study. *International Journal of Hydrogen Energy*, *59*, 465–479. https://doi.org/10.1016/j.ijhydene.2024.01.326

[169] Shahzad, M. W., Nguyen, V. H., Xu, B. B., Tariq, R., Imran, M., Ashraf, W. M., Ng, K. C., Jamil, M. A., Ijaz, A., & Sheikh, N. A. (2024). Machine learning assisted prediction of solar to liquid fuel production: a case study. *Process Safety and Environmental Protection*, *184*, 1119–1130. https://doi.org/10.1016/j.psep.2024.02.060

[170] Yi, J., Qi, Z., Li, X., Liu, H., & Zhou, W. (2024). Spatial correlation-based machine learning framework for evaluating shale gas production potential: A case study in southern Sichuan Basin, China. *Applied Energy*, *357*, 122483. https://doi.org/10.1016/j.apenergy.2023.122483

[171] Yang, Q., Ma, Z., Bai, L., Yuan, Q., Gou, F., Li, Y., Du, Z., Chen, Y., Liu, X., Yu, J., Zhou, X., Qian, C., Liu, Z., Tian, Z., Zhang, A., Sun, K., Ding, L., Tang, C., Meng, T., Min, F., & Zhou, Y. (2024). Machine learning assisted prediction for hydrogen production of advanced photovoltaic technologies. *DeCarbon*, *4*, 100050. https://doi.org/10.1016/j.decarb.2024.100050

[172] Elaziz, M. A., Zayed, M. E., Abdelfattah, H., Aseeri, A. O., Tag-Eldin, E. M., Fujii, M., & Elsheikh, A. H. (2024). Machine learning-aided modeling for predicting freshwater production of a membrane desalination system: A long-short-term memory coupled with election-based optimizer. *Alexandria Engineering Journal*, *86*, 690–703. https://doi.org/10.1016/j.aej.2023.12.012

[173] Leng, L., Zhou, J., Zhang, W., Chen, J., Wu, Z., Xu, D., Zhan, H., Yuan, X., Xu, Z., Peng, H., Yang, Z., & Li, H. (2024). Machine-learning-aided hydrochar production through hydrothermal carbonization of biomass by engineering operating parameters and/or biomass mixture recipes. *Energy*, *288*, 129854. https://doi.org/10.1016/j.energy.2023.129854

[174] Chen, H., Chen, J., & Ding, J. (2021). Data evaluation and enhancement for quality improvement of machine learning. *IEEE Transactions on Reliability*, *70*(2), 831–847. https://doi.org/10.1109/tr.2021.3070863

[175] Veiber, L., Allix, K., Arslan, Y., Bissyandé, T. F., & Klein, J. (2020). Challenges Towards Production-Ready Explainable Machine Learning. *2020 USENIX Conference on Operational Machine Learning (OpML 20)*, 21–23. https://www.usenix.org/conference/opml20/presentation/veiber

[176] Heck, P., Schouten, G., & Cruz, L. (2021). A Software Engineering perspective on building Production-Ready Machine Learning Systems. In V. Chkoniya (Ed.), *Handbook of Research on Applied Data Science and Artificial Intelligence in Business and Industry* (pp. 23–54). IGI Global Scientific Publishing. https://doi.org/10.4018/978-1-7998-6985-6.ch002.

Cite this Article

*Parankush Koul, and Yashavanth Siddaramu, "**MACHINE LEARNING IN PRODUCTION ENGINEERING: A COMPREHENSIVE REVIEW**", International Journal of Multidisciplinary Research in Arts, Science and Technology (IJMRAST), ISSN: 2584-0231, Volume 3, Issue 5, pp. 01-33, May 2025.*

Journal URL: *https://ijmrast.com/*

DOI: *https://doi.org/10.61778/ijmrast.v3i5.131*