UNIVERSITÄT SIEGEN

Analysis of Production Data and Predictive Analysis during the Production of Machine Tool

Master Thesis

Chair of Business Information Systems and New Media
submitted by

Maulik Ramani
Matr.-Nr.: 1650796

29.01.2025
at the Universität Siegen

First examiner:	Prof. Dr. Volker Wulf
Second examiner:	Herrn Hussain Abid Syed

Declaration of Honor

I hereby certify, Maulik Ramani, on my honor that I wrote this master's thesis with the title: "Analysis of Production Data and Predictive Analytics during the Production of Machine Tools" independently and without outside help and no others than used the tools specified. The parts of the work whose wording or meaning were taken from other works are always identified with the source indicated. The work has not yet been published or presented in any other form as an examination. I have understood the meaning of the honorary assurance and the consequences under examination law. I agree to the storage of my Master's thesis for the purpose of a plagiarism check. I confirm that the content of the electronic version is identical to the printed version.

Siegen, 29-01-2025

Place, Date

Signature

Abstract

In the production of roll grinding machine, a significant challenge lies in selecting the optimal processing steps and parameter configurations to achieve high-quality results. Small adjustments to parameters often lead to notable variations in output quality, making it difficult for operators to consistently find the "Right" parameters that optimize performance. This thesis leverages artificial intelligence (AI) to analyze extensive production data collected throughout the machining processes, with the goal of establishing a “Golden State” (an ideal set of steps and parameters for achieving optimal processing results; a defined state within the context of the use case company). In order to achieve this objective, initially, a structured SQL database, which stores input parameters, measurement results, and derived analysis data was setup, where i have automated the collection and evaluation of production data. Then by integrating predictive AI models (like LSTM Networks, Random Forest, GRU Networks), the optimal parameters were defined, curated and featured but also to provide real-time feedback during production, validating if the current process remains on the optimal path. The results yielded minimum error rate with respect to optimal grinding time, indicating the implications for the so-called golden state. This thesis contributes to the practice-centered use of machine data with AI algorithms in order to establish the boundaries of data-driven manufacturing, validated through the context of our use case company. This approach promises to improve decision making for machine tool production, reduce variability, and enhance product quality through data-driven predictive insights. In the later stages, the AI will further suggest parameter adjustments needed to realign the process with the "Golden State".

Keywords: Data Analysis, SQL Database, Data-driven Insights, Predictive Analytics

Abstrakt

Bei der Herstellung von Walzenschleifmaschinen liegt eine große Herausforderung in der Auswahl der optimalen Bearbeitungsschritte und Parameterkonfigurationen, um qualitativ hochwertige Ergebnisse zu erzielen. Kleine Anpassungen der Parameter führen oft zu erheblichen Schwankungen in der Ausgabequalität, so dass es für die Bediener schwierig ist, stets die "richtigen" Parameter zur Leistungsoptimierung zu finden. Diese Studie nutzt künstliche Intelligenz (KI), um umfangreiche Produktionsdaten zu analysieren, die während des gesamten Bearbeitungsprozesses gesammelt wurden, mit dem Ziel, einen "Goldenen Zustand" zu ermitteln (einen idealen Satz von Schritten und Parametern zur Erzielung optimaler Bearbeitungsergebnisse; ein definierter Zustand im Kontext des Anwendungsfalls Unternehmen). Um dieses Ziel zu erreichen, wurde zunächst eine strukturierte SQL-Datenbank eingerichtet, die Eingabeparameter, Messergebnisse und abgeleitete Analysedaten speichert, wobei ich die Sammlung und Auswertung von Produktionsdaten automatisiert habe. Dann wurden durch die Integration von prädiktiven KI-Modellen (wie LSTM-Netze, Random Forest, GRU-Netze) die optimalen Parameter definiert, kuratiert und dargestellt, aber auch um Echtzeit-Feedback während der Produktion zu liefern und zu überprüfen, ob der aktuelle Prozess auf dem optimalen Weg bleibt. Die Ergebnisse ergaben eine minimale Fehlerquote in Bezug auf die optimale Schleifzeit, was auf die Auswirkungen auf den sogenannten goldenen Zustand hinweist. Diese Arbeit leistet einen Beitrag zur praxisorientierten Nutzung von Maschinendaten mit KI-Algorithmen, um die Grenzen der datengesteuerten Fertigung festzulegen, validiert durch den Kontext unseres Anwendungsfalls Unternehmen. Dieser Ansatz verspricht, die Entscheidungsfindung in der Werkzeugmaschinenproduktion zu verbessern, die Variabilität zu verringern und die Produktqualität durch datengestützte Vorhersagen zu erhöhen. In späteren Phasen wird die KI weitere Parameteranpassungen vorschlagen, die erforderlich sind, um den Prozess auf den "Golden State" auszurichten.

Schlüsselwörter: Datenanalyse, SQL-Datenbank, Datengesteuerte Einblicke, Prädiktive Analyse

Index

Illustration Index

Table Index

Abbreviation

GDP	*Gross Domestic Product*
IoT	*Internet of Things*
CNC	*Computer Numerical Control*
SCADA	*Supervisory Control and Data Acquisition*
MES	*Manufacturing Execution Systems*
DCS	*Distributed Control Systems*
PLCs	*Programmable Logic Controllers*
SPC	*Statistical process control*
AI	*Artificial Intelligence*
ML	*Machine Learning*
ADS	*Automation Device Specification*
NC	*Numerical Control*
PA	*Predictive Analytics*
AR	*Autoregressive*
ARIMA	*Autoregressive Integrated Moving Average*
MA	*Moving Average*
MAPE	*Mean Absolute Percentage Error*
RMSE	*Root Mean Squared Error*
LDA	*Linear Discriminant Analysis*
LSTM	*Long Short-Term Memory*
GRU	*Gated Recurrent Unit*
RNN	*Recurrent Neural Network*
LightGBM	*Light Gradient Boosting Machine*
SVM	*Support Vector Machines*
KNN	*K-Nearest Neighbors*
SFS	*Sequential Forward Selection*
MI	*Mutual Information*
PCA	*Principal Component Analysis*

MSE	*Mean Squared Error*
MAE	*Mean Absolute Error*
TP	*True Positives*
TN	*True Negatives*
FP	*False Positives*
FN	*False Negatives*
RL	*Reinforcement Learning*
t-SNE	*t-Distributed Stochastic Neighbor Embedding*
RF	*Random Forest*
GBR	*Gradient Boosting Regressor*
XGBoost	*Extreme Gradient Boosting*
SVR	*Support Vector Regression*
CatBoost	*Categorical Boosting*

1. Introduction

The development of the manufacturing sector has impacted global economic expansion and technological breakthroughs to a great extent [1]. Manufacturing maintains a vast supply chain, which crosses sectors such as automotive, aerospace, electronics, energy, and heavy industries, fosters innovation, and provides necessities. This industry employs various cutting, molding, grinding, and finishing techniques to convert raw materials into high-precision parts and components, with the help machine tools. Machine tools are specialized equipment used in shaping and finishing metals or other hard materials by machining, grinding, or cutting operations. Machine tools are accorded a special place within the manufacturing industry because they help in keeping up the quality, accuracy, and scalability needed for contemporary production processes apart from rendering it physically possible to create items.

In the past few decades, manufacturing has gone through a number of changes, starting from the mass production in the middle of the 20th century to lean manufacturing and, most recently, to automated and data-driven systems, or Industry 4.0 [2]. Under this new paradigm, manufacturing systems have evolved into highly automated, networked, and data-driven systems. Since machine tools are more than just physical tools utilized in production, they play a very important role in this transformation. They also generate useful information about operational performance, which can be exploited to improve manufacturing outcomes [3].

1.1 Context and Importance of Machine Tool Production

Machine tools are important to manufacturing not only because of their operation capabilities but also because they can produce crucial data in the analysis and improvement of processes [4]. The manufacturing firms face continuous pressure to supply a high-quality component at low cost and within a strict timeline due to rising global competition. These requirements demand optimal set-ups and stability in operations but more importantly, efficiency in the operation of machine tools. Thus, machine tools are needed to produce consistent, high-quality work that will meet the competitive requirements in modern manufacturing.

1.1.1 Overview of the Manufacturing Industry

Machine tools are essential to the production of highly precise crucial parts in industries including consumer electronics, automotive, aerospace, and heavy machinery [2]. Their ability to manage a wide variety of materials and geometries makes them a key part of manufacturing processes that require both precision and flexibility. For example, the automotive industry relies on machine tools for the manufacture of chassis systems, engine components, and transmission components, all of which must meet very strict requirements concerning durability and tolerances. The same is true when creating high-performance aircraft components, the aerospace industry requires machine tools capable of handling light yet strong materials such as titanium and composites.

With industrial centers scattered across the globe, in North America, Europe, and Asia, manufacturing contributes greatly to both the global Gross Domestic Product (GDP) and employment [5]. Changes and improvements in one sector will affect the norms and practices of manufacturing in another, hence showing the linkages of this global network. The industry has adopted technologies supporting more agile and intelligent production processes over the past decade, driven by increasing complexity in manufactured goods and consumer desire for customization. This change has also increased the need for machine tools that can handle complicated jobs, adjust to particular production needs, and guarantee the accuracy and uniformity required by sectors that depend on highly regulated operations.

1.1.2 Importance of Efficient Machine Tool Production

Machine tool production, to be competitive in the industrial sector, has to be effective. Accuracy, consistency, adaptability, and the possibility of reducing waste and energy consumption together form the efficiency of machine tools, transcending just simple operational speed or power [6]. The efficiency-oriented machine tools can cut down manufacturing time and costs drastically, eventually improving the quality and uniformity of the final output. In addition, efficient machine tools provide regular output and enable the firms to meet massive production levels through reduced time in production delays and downtime.

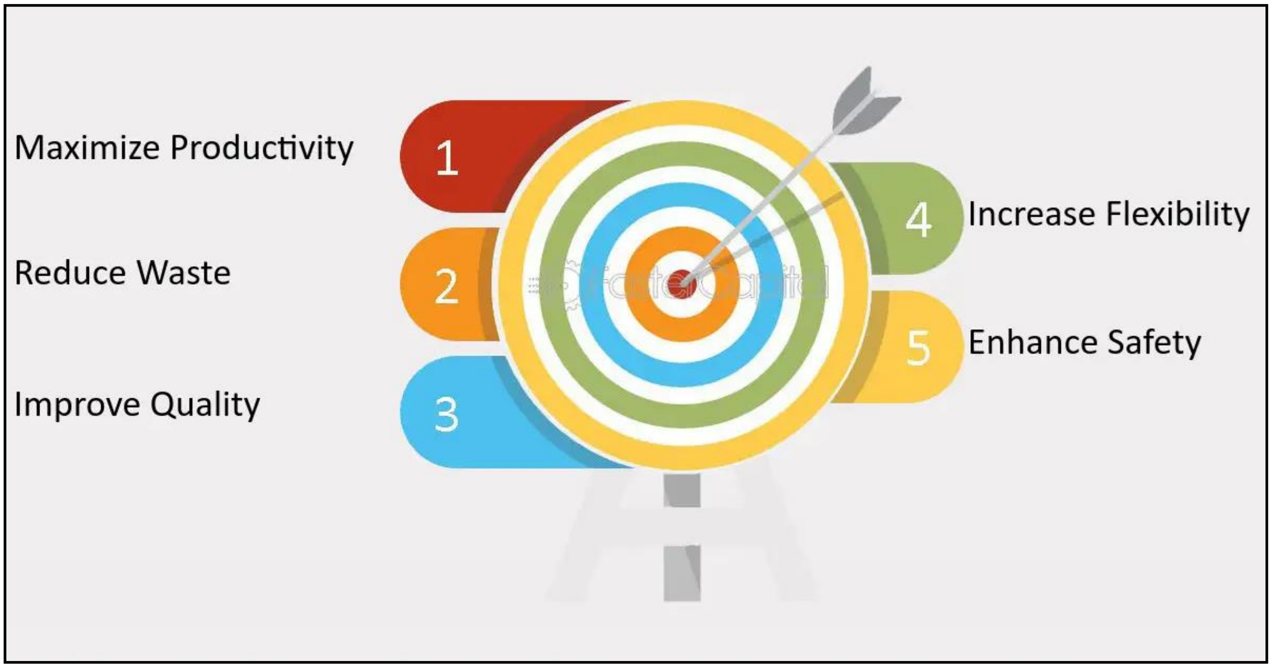

Figure 1: The Importance of Efficiency in Manufacturing [63]

Nowadays, increases in digital technology directly affect the efficiency of machine tools. Real-time information on machine performance, including wear and tear, energy consumption, and operational abnormalities, can be known by manufacturers using a combination of sensors, Internet of Things (IoT) capabilities, and data analytics [1]. This then goes straight to predictive maintenance, essential for the minimization of unplanned downtimes through the use of real-time data to predict machine tool problems before they occur. And, at the same time, thanks to data analytics, adaptive control systems can automatically adjust machine tool parameters according to real-time production demands in order to deliver high levels of customization while preserving accuracy.

In modern manufacturing, the more important role that machine tool production plays in the satisfaction of changing needs increases its importance [7]. Quality and safety, which cannot be compromised in industries like aerospace and healthcare device manufacture, depend in large part on precision machining and high-quality production. Effective machine tools also support green production methods as they help in waste reduction and energy consumption with the shift of manufacturing to more green-friendly methods.

1.4 Problem Statement

The manufacturing industry continuously faces high demands for productivity, precision, and quality with minimal wastes and operational costs [15]. Manufacturers who deal in machine tools, especially grinding machines, have to continually improve processes of production by concentrating on core competencies in order to remain competitive and meet the high demand for better quality products at reduced cost. These variations in parameters such as feed rate, wheel speed, depth of cut, and material properties greatly influence the efficiency of the operation and the accuracy of the results during grinding.

Minimizing errors to achieve the ideal grinding time has always been one of the primary challenges faced in machine tool production, particularly in grinding operations. Most complex relationships between feed rate, wheel speed, wheel type, pressure, depth of cut, and material qualities are brought into grinding operations, which easily cause production cycle abnormalities and inefficiencies. These differences are often responsible for expensive delays in production, increasingly high rates of error, and low-quality products. Even now, technological advancement has not allowed the manufacturers to predict exactly the optimal operating conditions necessary for fewer errors and more productions. A proper understanding of the variables affecting grinding efficiency and a means to estimate optimal operating conditions is essential in order to effectively sort out these problems.

There is a growing interest on the part of manufacturers in data-driven solutions for informed decision-making on operations in light of these challenges [16]. Traditional methods of grinding are too reliant on operator know how and trial and error techniques, which cannot meet the challenge imposed by high precision and efficiency. These methods may also be very time-consuming and may result in variable results given the various operators' experience and skill level. The result has been a strong impetus towards the adoption of the data-centric method, which goes on to provide predictive insights that look into the grinding operation outcomes in a manner that is consistent, repeatable, and very precise indeed. By proactive predictions and adjustments of the grinding parameters, manufacturers can make closer attainment to consistent product quality, reduced cycle times, and reduced operational cost.

Recent development within the field of Artificial Intelligence (AI), Machine Learning (ML), and IoT has opened new areas for manufacturing innovation [17]. Today's manufacturers, using IoT-enabled devices and sensors, have unparalleled capability in collecting real-time data on the condition, operating parameters, and product quality. These algorithms have the ability to analyze volumes of data in a quest to learn from them and process the same for intricate patterns and connections that could not be deciphered earlier. Technological advances are making it possible for manufacturers to transition from traditional methods to data-centric methods that give much better control over grinding parameters and, therefore, productivity [16]. By integrating these advanced tools with grinding processes, manufacturers can forecast optimal settings for several parameters, predict outcomes based on real-time data, and make dynamic adjustments in operations to secure the best results.

The motivation for this research is to turn these technological advances towards solving longstanding problems that challenge grinding operations. Based on historic production data, this work proposes predictive models that furnish actionable insights with which a manufacturer could proactively decide on improving their overall cost-effectiveness in grinding operations by reducing error rates and improving the efficiency of the process. Successfully deploying such predictive models would not only help bring production efficiency and product quality into a different league, but it would also put the manufacturer in a better position to meet the demands of a fast-moving, quality conscious world market.

1.5 Research Questions

The following important research questions will be addressed in order to fulfill the stated objectives:

1. Which particular parameters in the grinding process most significantly affect overall grinding efficiency?

- Identify and analyze the key operational parameters, which influence both the accuracy of the grinding process, and the time required for operations.

2. How can predictive analytics be utilized to develop a model that predict the optimal grinding parameters needed to minimize errors and grinding time?

- The methodologies and techniques for using historical production data to build a predictive model that suggests the best parameters for reducing grinding errors and optimizing time.

3. What are the impacts of implementing data-driven predictive models on the grinding process?

- Evaluate the practical outcomes of applying predictive analytics in grinding operations that result from optimized grinding parameters.

2. Literature Review

2.1 Role of Production Data in Modern Manufacturing

Modern manufacturing demands production data in order to increase productivity, maintain product quality, and drive innovation [3]. Production data is a large set of operational, performance, and quality measurements taken at various stages in the process of manufacturing. This can be used by manufacturers to move away from traditional reactive approaches towards proactive and predictive methods, optimizing output and avoiding downtime. Data-driven manufacturing aligns with Industry 4.0 principles, which emphasize the integration of digital technologies such as IoT, big data analytics, and artificial intelligence in creating connected, intelligent production environments. The role of production data within different manufacturing processes and the technologies applied to collect, process, and use efficiently the generated data.

2.1.1 Production Data Across Various Manufacturing Processes

There is a lot of production data generated, and there are many uses for that information: [8] it could be used in estimating equipment health, tracking how effective a process is, and ensuring the quality of a product. Different forms of data are involved in different manufacturing processes, all of them are important in helping control the process, reducing variability, and ensuring production standards. For instance:

Machining Processes: Data such as spindle speed, feed rate, temperature, and vibration are critical in monitoring machine tool health and ensuring precision. This data helps prevent tool wear and ensures that each component meets strict dimensional tolerances.
Welding Processes: In welding, production data includes voltage, current, speed, and temperature, all of which affect the strength and quality of the weld. Data monitoring enables real-time adjustments to improve weld quality and reduce defects.

Manufacturers can take holistic data-driven approaches in integration with equipment performance, product quality, and process efficiency by systematically collecting and analyzing data from these myriad processes. This will provide a solid basis for process optimization and predictive maintenance.

In the manufacturing environment, there are many specialized, interconnected systems that record, track, and analyze production information. These may include sensor networks, control systems, CNC systems, SCADA, MES, and Internet of Things-based remote monitoring tools [9]. The different data types and monitoring capabilities provided by each system allow comprehensive visibility and control of production.

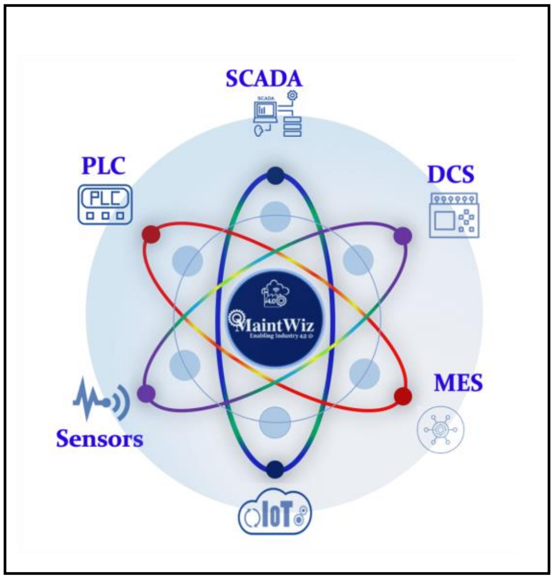

Figure 2: Sensor Networks, Control Systems, CNC Systems, SCADA, MES, and Internet of Things-based for Record, Track, and Analyze Production Information [12]

2.1.1.1 CNC System Integration: Capturing Operational Data

Computer Numerical Control systems (CNC) are an integral part of automated production, particularly where machine tools are concerned [9]. They translate preprogrammed instructions for machining operations into precise movements of the tool. They also record various vital operating data regarding the position of the tool, rapidity, feed, and power consumption. This data is very essential in monitoring machine performance and pinpointing variances in operation that could lead to flaws or wear on the equipment. Thus, CNC systems form the basis of closed-loop control. this is where dynamic adjustment becomes possible through feedback in real-time data, minimizing errors and thereby assuring high precision throughout repetitive activities.

2.1.1.2 SCADA and MES Systems: Plant-Wide Data Monitoring

Plant-wide monitoring and control capabilities are provided using Manufacturing Execution Systems (MES) and Supervisory Control and Data Acquisition (SCADA) [12]. SCADA systems centrally monitor production parameters, such as temperature, pressure and flow rates by collecting real-time data from sensors like cabinets for the cooling unit in relation to current IT hardware power values or other instruments placed at various locations around a factory. MES, in comparison, is able to generate a complete view of manufacturing activities by blending that real-time data with production planning and scheduling information as well as resource allocation and quality control. Combining the information from control systems and SCADA, MES is used to improve schedule accuracy between planned schedule execution including batches cut by operators versus plan.

2.1.1.3 Sensor Systems: Real-Time Condition Monitoring

Sensors are key to real-time condition monitoring in manufacturing [12]. They are embedded in the equipment to measure temperature, vibration, force and humidity. These metrics are critical to equipment health as real-time monitoring can detect early signs of component wear or irregularities that can lead to equipment failure. Condition based maintenance with sensor data means manufacturers can fix problems before they become downtime. And advanced sensors are crucial in quality assurance as they monitor the environment in which sensitive components are manufactured to ensure they are within optimal ranges.

2.1.1.4 Control Systems: Managing Production Processes

From controlling machine functions to multi-step processes, control systems are used to manage and automate many industrial processes. Distributed Control Systems (DCS) and Programmable Logic Controllers (PLCs) are used in manufacturing to sequence production stages and control machines. Manufacturers can monitor process variables like temperature and pressure using data from control systems to ensure consistency and compliance to standards. These systems can also change process parameters in real-time based on data to maintain optimal operating conditions even when production conditions or material change. So, control systems give quick adaptability to changing production demands by increasing production flexibility as well as process stability.

2.1.1.5 Data Export and Analysis: Enhancing Process Optimization

One of the key parts of industrial optimization and continuous improvement is exporting and analyzing production data. Data analysis can show you process bottlenecks, hidden trends and where you can increase efficiency. Production equipment and control systems output data to central data platforms where advanced analytics tools use machine learning and statistical methods to find patterns and anomalies. Predictive analytics for example can predict when equipment needs maintenance based on historical data and reduce downtime and increase asset life. Production data analysis also enables root cause analysis which helps you find and fix the underlying causes of process inefficiencies and quality defects.

2.1.1.6 IoT Capabilities: Remote Monitoring

With IoT you can monitor and control manufacturing equipment remotely, which means production data gets a boost [9]. Sensors, machines and control systems are connected via IoT networks which send real time data to cloud based platforms. With this you can monitor production environments remotely, so response times to equipment failures and production downtime is faster. By combining data from multiple units IoT enables predictive maintenance by allowing centralized analysis to find issues before they become problems. Remote monitoring of machines and processes means you can be more responsive, more adaptable and ensure business continuity.

2.2 Key Benefits of Production Data Analysis

Analyzing production data offers many benefits that are relevant to various areas of manufacturing, particularly in the production of machine tools. By leveraging data-driven insights, manufacturers can significantly reduce costs, enhance operational efficiency, and make better-informed decisions. This section delves into the key advantages of production data analysis, focusing on its applications in machine tool manufacturing and the positive impacts on cost savings and production efficiency.

2.2.1 Applications of Production Data Analysis in Machine Tool Production

The analysis of production data plays a vital role in various machine tool manufacturing sectors, helping manufacturers enhance efficiency and ensure high-quality outcomes. Some of the most notable applications include:

Predictive Maintenance [10]**:** Predictive maintenance helps in the anticipation of equipment breakdowns before they actually happen. It enables manufacturers to view trends pointing to imminent machine failures by analyzing past operational data, which may include vibration, temperature, and runtime measurements. This, therefore, allows for scheduled interventions that help reduce maintenance costs and minimize unplanned downtime.

Process Optimization [10]**:** Data analytics could be used on machining operations in order to detect inefficiencies or bottlenecks in the manufacturing process. Operators will be able to monitor key performance indicators, including cycle time, throughput, and tool wear rates, by applying Statistical process control (SPC) approaches. This allows them to make well informed modifications to improve overall process efficiency.

Quality Control [10]**:** Data analysis helps in quality control by identifying the patterns and any departure from the set standards of quality. The manufacturer will, therefore, be able to identify the causes for defects, put corrective measures in place, and eventually scrap rates shall decrease based on analyses from in-process inspections and final product testing. Rework costs can be minimized, and at the same time, consistency in quality standards is maintained with data-driven insights.

Figure 3: Use cases of Predictive Models in the Manufacturing Sector [11]

Production Planning and Scheduling [11]**:** Advanced data analytics technologies make it possible to have a more accurate production planning and scheduling consideration of previous performance data, order trends, and machine availability. Through the use of these tools, manufacturers cut down lead times, increase flexibility within production schedules, and improve resource utilization. Hence, production processes become well prepared for adjustments in changing demand.

Performance Benchmarking [11]**:** The manufacturer will be able to compare his performance with internal metrics or industry standards. This can be done using production data analysis. Organizations will be able to find best practices, exchange information, and put strategies into action that implement improvements in overall performance by comparing operational data from various equipment or production lines.

Resource Utilization [12]**:** Manufacturers can evaluate labor, material, and equipment utilization rates by conducting an efficient study of production data. Businesses can take focused steps to improve their resource allocation and boost productivity and cut waste by identifying areas where resources may be excess or underutilized.

These solutions allow businesses to use insights for ongoing improvement and long-term competitive advantage, making production data analysis a fundamental pillar for attaining operational excellence in machine tool production.

2.2.2 Impact of Production Data Analysis on Production Efficiency and Cost Savings

Analysis of production data contributes considerably toward cost reduction and efficiency in production, radically changing the way factories work. The highest benefits of effective data analysis across industries are enumerated in the coming points:

Enhanced Operational Efficiency [14]**:** With production data analytics, manufacturers can achieve better operational efficiency by shortening cycle time, optimizing machine utilization, and streamlining processes. This allows for real-time performance monitoring to track inefficiencies quickly and take rapid corrective action to improve workflows and output.

Reduced Downtime: Data-driven predictive maintenance techniques have helped to reduce unscheduled downtime by a large margin. Manufacturers can sustain extended production flows by predicting equipment breakdowns and thus planning repair during downtime. This reduces overall production time since machines are available when they are actually needed.

Lower Production Costs [14]**:** Data analysis enables manufacturers to reduce production costs by cutting down on waste, rework, and ensuring resources are used optimally. This can help manufacturers lower their costs per unit of production while preserving or even raising quality by finding out about inefficiencies and making the quality control stronger. The financial benefits increase profitability.

Improved Quality Assurance [8]**:** Manufacturers will be able to have strict quality standards by analyzing the manufacturing data on a timely basis. Organizations can enhance their quality assurance practices and reduce recalls, rework cases, and warranty claims by finding and reducing the variables causing product faults. This would not only reduce expenses related to defects but also improve client loyalty and satisfaction.

Data-Driven Decision Making [8]**:** Instead of intuition or tradition, producers can make the best possible decision with the knowledge gained from production data analysis. As part of this new change toward a data-centric culture, the organization will be able to react more nimbly to marketplace changes, optimize production processes, and closely align operational goals to business objectives.

Sustainable Practices [14]**:** The production data analysis helps to identify areas where material and energy use could be optimized, thus promoting sustainable practices of manufacturing. Manufacturers may reduce operating costs and play their role in maintaining environmental sustainability by minimizing waste and engaging energy-efficient procedures through monitoring production data with care.

2.3 Production Data in Roll Grinding Machines

Manufacturers who want to monitor and enhance the efficiency and quality of their operations need to apply production data precisely in the context of roll grinding operations [18]. A very important component in many industries, roll grinding comprises the accurate grinding of cylindrical rolls used in specific applications that require machinery parts with smooth and even surfaces. The integration of vast amounts of data gathered from grinding machines has revolutionized the way roll grinding process is conducted using data science and modern industrial technologies. This information allows for the understanding of machine efficiency, process variables, product quality, and it encourages developments in predictive and preventive maintenance and better management of production processes.

With the use of sophisticated sensors and control systems, roll grinding machines are able to record a big range of data, including feed rate, spindle speed, temperature, vibration levels, and results data in real time [17]. Operators can increase grinding precision, improve the life of the equipment, and lower operating costs by gathering, evaluating, and interpreting this data. This section examines how roll grinding machines gather data and how operational data might improve roll grinding procedures.

2.3.1 Overview of Data Collection in Roll Grinding Machine

Data gathering is a multi-tier architecture in which real-time data from the operations of the machine are collected and stored in database for later analysis. The major components involved in the system are advanced sensors, communication protocols, control software, and data storage solutions.

Machine Integrated Sensors [19]: Sensors are basically the backbone of any roll grinding machine for data acquisition. These sensors are mounted on the main spindle, grinding wheel, and other important machine components, monitoring key process variables such as temperature, speed, torque, and vibration. Temperature sensors measure the produced heat at the workpiece and at the grinding wheel, which may influence surface finish and tool wear. Torque sensors measure the force applied by the grinding wheel. It serves as an indicator of power efficiency of the machine and load on the tool. Abnormal vibrations, which may be a sign of imbalance or wear in the grinding components, can be detected with the aid of vibration sensors. EtherCAT, an industrial Ethernet technology renowned for its fast, deterministic data transfer capabilities, is the communication protocol used to convey the data from these sensors. All monitored characteristics are instantly available for analysis and control because to EtherCAT's ability to facilitate real-time communication between sensors and the machine's control systems.

Drives and EtherCAT Communication [20]: EtherCAT network creates a synchronized environment whereby all the aspects of machine operation are continuously monitored and altered by the integration of drives besides connecting sensors, including other control components to create powered motors that drive machine movements. With the high bandwidth of EtherCAT, data from numerous sensors can be delivered to the central processing units in near real time to ensure that changes in machine conditions are timely detected and managed. Besides this, the communication architecture also supports the ADS interface, which stands for Automation Device Specification and supports data interchange across different layers of machine control systems so that sensor and drive data may effectively reach the TwinCAT core processing and control software.

Centralized Control with TwinCAT [20]: The main software platform for control and data processing in roll grinding machines is called TwinCAT (The Windows Control and Automation Technology). Numerical Control (NC) and Programmable Logic Controller (PLC) are its two primary components, which work together to convert sensor data into machine commands that can be implemented. By precisely controlling the machine's movements, the NC component makes sure that the grinding wheel travels in the right direction and maintains the necessary feed rate. In roll grinding, where accuracy and surface polish are crucial, this precision is important. In order to ensure safe and effective operation, the PLC component coordinates the actions of many machine components and controls the overall machine logic. The PLC implements safety procedures and modifies operational parameters as necessary by interpreting real-time input from sensors and drives. A logic controller in TwinCAT analyzes incoming sensor data and decides what has to be changed to maintain the machine in ideal operating conditions. This includes adapting to variations in torque or temperature that could otherwise impact grinding quality or lead to premature wear on machine components.

ADS (Automation Device Specification) for Data Communication [21]: It is developed for data exchange between the different software modules, for instance the communication between the NC and the PLC. It serves as a link between the TwinCAT system and other systems, such as the history database and KPNT software. ADS allow data to be incorporated into a broader manufacturing execution system (MES) for plant-wide monitoring or sent to external analysis tools. This makes it possible for the roll grinding machine's operational data to support more comprehensive production insights and optimization plans inside the manufacturing facility.

Data Transfer and Service Layer (KPNT Software) [21]: Through the ADS interface, the TwinCAT-processed data is transmitted to the KPNT software and related service module. In addition to providing a user interface for monitoring and troubleshooting, this layer is essential for managing the machine's operational data. As a middleware, KPNT software links with other plant systems to enable integrated production monitoring and arranges data for operators' easy access. Maintenance personnel can evaluate machine conditions and address possible problems thanks to the diagnostics capabilities of KPNT's service module. It is simpler to identify patterns, streamline grinding procedures, and apply enhancements based on insights from operational data when operators have access to both real-time and historical data through this software.

Historical Data Storage (History Database): All information gathered from roll grinding operations is permanently stored in the History Database. This comprises processed and real-time data that is systematically saved for later examination. Historical information about machine parameters, grinding outcomes, tool wear, and maintenance events is stored in the database. Trend analysis, which uses data trends to provide insights into component wear rates, production efficiency, and product quality over time, is made possible by this collection. By enabling maintenance personnel to foresee when particular machine parts may break down or need to be serviced, stored data also makes predictive maintenance possible. For instance, a steady upward trend in torque data may point to a growing problem with the spindle or grinding wheel, requiring preventive repair to avoid unplanned downtime.

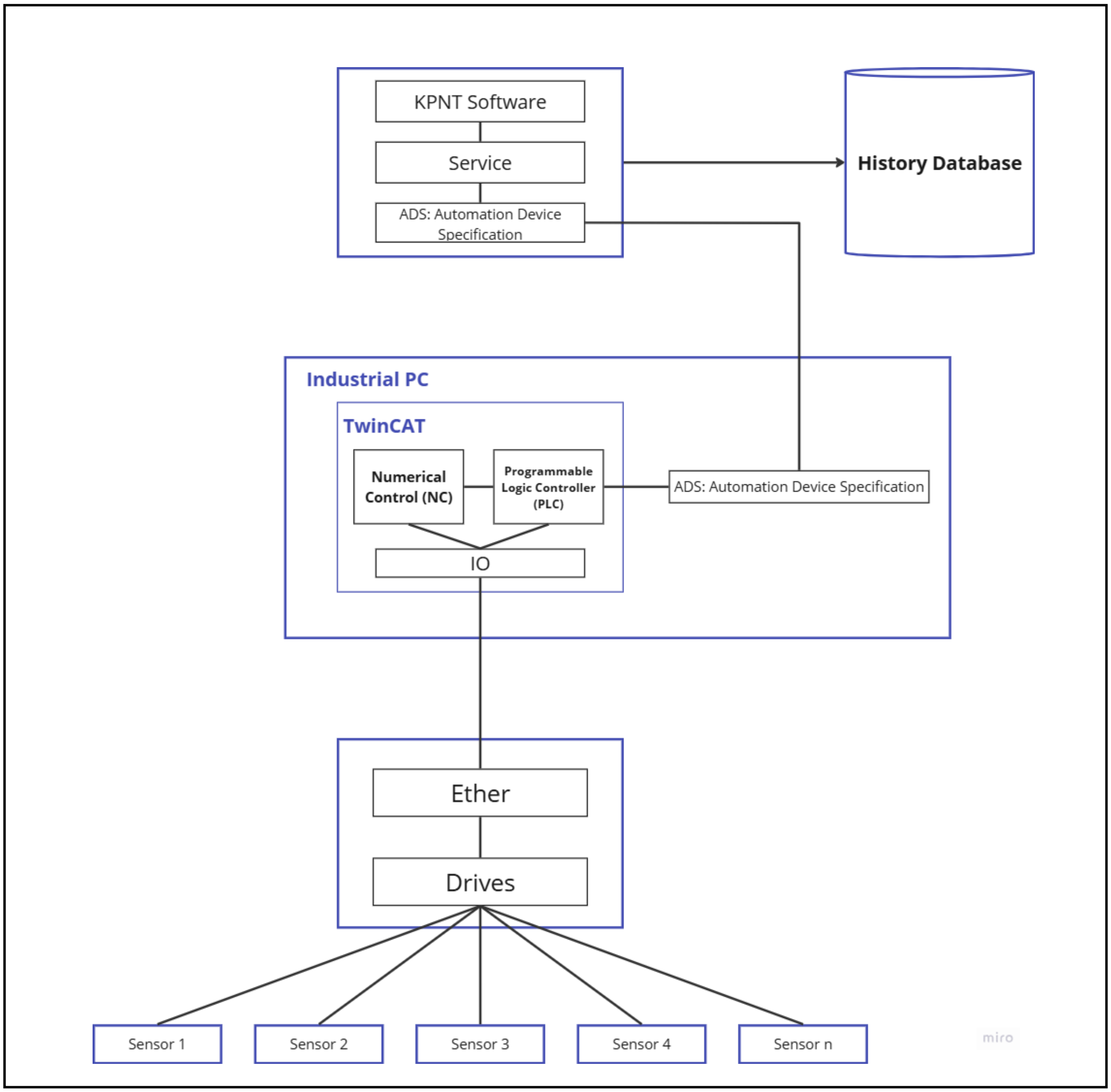

Figure 4: Flowchart of Data Collection in Roll Grinding Machine

A high degree of accuracy and efficiency in data collecting can be attained by roll grinding machines through the integration of sensors, control systems, condition monitoring, and data logging within a structured framework. Improved process management and maintenance are made possible by this data, which facilitates both retrospective analysis and real-time monitoring.

2.3.2 Importance of Operational Data in Roll Grinding

For roll grinding processes to be optimized and efficient, operational data is essential. Achieving consistent product quality, increasing process efficiency, and enabling proactive maintenance strategies all depend on the gathering and analysis of operational data in contemporary manufacturing environments, especially in sectors that depend on high-precision roll finishing. Throughout the grinding process, producers may improve decision-making, prolong equipment lifespans, and lower total operating costs by methodically collecting and analyzing data. Several crucial elements help to clarify the significance of operational data in roll grinding:

Process Optimization [10]**:** Throughout production, operators can dynamically monitor and modify grinding parameters thanks to the ongoing collection of real-time operating data. For instance, operators can instantly modify important parameters like feed rates or grinding wheel speed by utilizing sensor feedback, such as force or temperature readings. By preventing possible harm to the roll and the grinding machine, this degree of real-time control makes sure that the process stays within ideal operating parameters. Errors are greatly decreased, rework is reduced, and more precise, superior results are produced when the grinding process can be adjusted on the run. Additionally, this ongoing monitoring makes it possible to spot process variations and take prompt corrective action before flaws affect the quality of the final product.

Predictive Maintenance [22]**:** Implementing predictive maintenance techniques, which are critical in averting unplanned equipment failures, also need operational data. Roll grinding machines can identify early indications of wear and tear or possible problems in important components by using integrated sensors to monitor characteristics including vibration, temperature, and noise emissions. Logs of historical data make it possible to spot patterns and trends that can point to new problems like imbalance, misalignment, or excessive tool wear. For example, alterations in vibration patterns may suggest that a spindle overhaul or bearing replacement is necessary. By addressing possible problems before they result in catastrophic failures, predictive maintenance which is fueled by constant condition monitoring helps manufacturers minimize unscheduled downtime and increase the machinery's operational lifespan.

Quality Control and Consistency [12]**:** It's critical to ensure consistent product quality in industries that need precision finished rolls. In order to guarantee the correct surface smoothness and dimensional accuracy of the rollers, roll grinding procedures frequently function under strict tolerance restrictions. A thorough record of every grinding operation is provided by operational data collected from machine sensors and control systems, guaranteeing that every roll satisfies the required specifications. Operators may make sure that the grinding process complies with the set quality criteria by monitoring real-time data, such as temperature, grinding wheel speed, and applied pressure. By doing this, operational data turns into a crucial instrument for preserving product uniformity across production cycles and guaranteeing that every roll satisfies the strict requirements needed for subsequent procedures like manufacturing or finishing.

Data-Driven Decision Making [18]**:** With access to wide ranging, real-time operational data, companies can migrate from reactive problem-solving to a more proactive, data-driven approach to decision-making. The comprehensive production data logs provide insightful information on long-term patterns and operational effectiveness. For instance, operators may better anticipate when tools will need to be replaced by examining previous data on tool wear rates. This helps them prevent needless downtime or the danger of utilizing worn-out equipment. In order to maximize grinding cycles and improve product quality, data-driven insights can also help guide decisions about how to modify grinding parameters like feed rates, depth of cut, or wheel speed. Manufacturers can increase overall production efficiency, minimize tool consumption, and optimize resource allocation by basing operational decisions on comprehensive performance data rather than conjecture.

Integration with Advanced Analytics [23]**:** The use of machine learning algorithms and advanced analytics in roll grinding is becoming more and more important as the technology advancement.

Manufacturers can use sophisticated analytics tools to extract relevant insights from the wide amounts of operational data gathered during each grinding process that would otherwise be challenging to find. To forecast the optimal grinding parameters for a certain material type or product requirement, for example, machine learning models can be taught to identify trends in the data. By continuously adjusting operational parameters, these predictive capabilities improve the grinding process and guarantee that every grind is tailored for optimal quality and efficiency. Predictive analytics can also identify maintenance requirements, which lowers the chance of unplanned stoppages and enables more accurate scheduling of machine downtime.

2.4 Types of Production Data in Roll Grinding

In roll grinding operations, effective production management is underpinned by the systematic collection, analysis, and application of various types of production data. This data provides important insights into crucial components of the grinding process, such as machine performance, product quality, and equipment health. Manufacturers may support proactive maintenance initiatives, increase productivity, and improve process efficiency by using production data to drive their decisions. Machine Parameters Data, Quality and Performance Data, and Maintenance and Condition Monitoring Data are the three primary categories into which roll grinding production data can be broadly separated. Each of these categories has a particular purpose, helping to enhance operating efficiency, maintain uniform product standards, and assure long-term equipment reliability.

2.4.1 Machine Parameters Data

The machine parameters data will include a wide range of variables measured directly from the operating roll grinding machine. Core operating parameters, along with extra measurements for the specific characteristics of the particular roll being serviced, compose this variety of data. These combined variables provide operators the capability to monitor and regulate the grinding process in real-time with the help of each roll to meet the required specifications and quality standards.

One of the most important parameters of a machine is the **feed rate**, which determines how quickly the grinding wheel moves across the surface of the roll [23]. Surface finish, grinding accuracy, and material removal rates can all be greatly impacted by changes in the feed rate. While a quicker feed rate might speed up the grinding process, it may also raise the chance of errors if improperly regulated. For example, a slower feed rate may result in a smoother finish but at the expense of longer processing periods. Likewise, the **speed of the grinding wheel** is important since it influences both the quality of the surface finish and the effectiveness of material removal. While lower speeds could reduce productivity, higher speeds can improve cutting efficiency but can cause surface burn if used excessively.

Another critical parameter is the **depth of cut** or the thickness of the material layer that is removed in one Stroke [24]. This variable does not only affect surface finish but also has an effect on machine loading and stability. Increasing the depth of cut can speed up material removal. however, it may also introduce vibrations or wear on the tool, which would decrease precision. On the other hand, shallower cuts may improve the precision but require more passes, thus elongating the grinding time [23]. **Coolant flow rate** and **temperature** are two other important parameters controlling the thermal conditions of grinding. Application of coolants relieves overheating and reduces thermal distortion, preserving the integrity of both roll and grinding wheel. A constant flow rate of coolant and optimum temperature reduce thermal stress in the roll and machine to help maintain dimensional accuracy and surface quality.

Besides these critical machine parameters, the physical forces applied, namely, **force** and **torque**, are also monitored to ensure that the grinding forces are within safe limits. These parameters give insight into the physical dynamics of the grinding process. Excessive force or torque may cause excessive wear on the grinding wheel, damage the roll, or create unstable grinding conditions. The observation of these forces enables the system to alert operators or automatically adjust the settings if the limits are exceeded, therefore protecting the machinery and workpiece from damage.

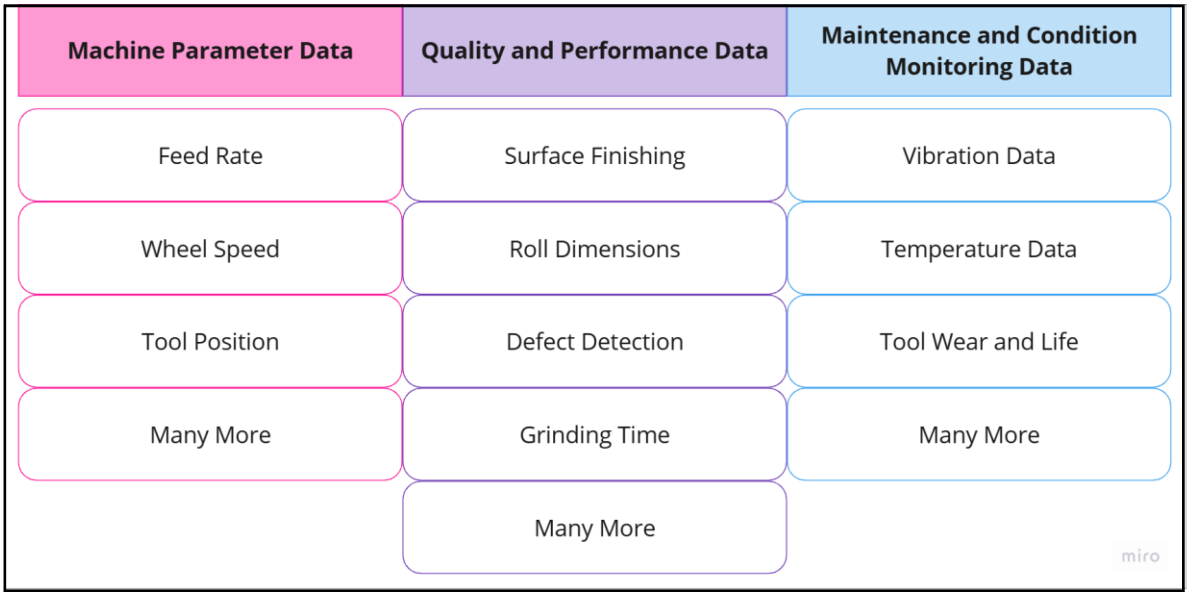

Figure 5: Production Data in Roll Grinding

Machine parameters in roll grinding will also involve detailed information on some roll specific characteristics, such as length, diameter, surface roughness before grinding, and crack detection [23]. **Length** and **diameter** measurements are basic for determining the dimensions of the roll and ensuring that the roll meets the specifications before grinding. Deviation in any of these dimensions may require adjustment in the grinding setup to arrive at the desired final dimensions. **Surface roughness** before grinding is another important measurement since it establishes a base to understand the effectiveness of the grinding action. Thus, post-grinding roughness values can be compared to the pre-grinding values to ensure the grinding has achieved the desired quality of **surface finish** [25]. Not least important, **crack detection** data provides information concerning the existence of cracks in the roll surface or immediately subsurface that can be initiated or aggravated by the grinding stresses. Early detection of cracks in these items allows operators to make informed decisions about the treatability of the roll, whether it can be processed or needs further treatment.

This will enable operators to keep the process within optimal operating ranges by continuously monitoring these machine parameters [24]. In such a way, stability in the grinding environment is guaranteed. Where any parameter exceeds beyond threshold values set, the system may send alerts or even automatically adjust the settings to avoid potential damage or quality issues. This is the extent of control possible with real-time Machine Parameters Data, which helps in securing better accuracy and stability in the grinding process, minimizes defects in the lot, and hence makes every roll much closer to optimal quality standards. Basically, Machine Parameters Data forms the backbone of a modern roll grinding operation for the very reason that it lets data-driven decisions be taken in order to optimize productivity, quality, and equipment health.

2.4.2 Quality and Performance Data

Quality and Performance Data is a small but critical set of metrics on the quality assessment of the finished roll and also effectiveness in the grinding operation [26]. These key data show whether the grinding process meets standard requirements for effectiveness in precision and productivity. Continuous monitoring and analysis of such metrics allow the operators to ensure that the grinding process is producing high quality rolls constantly and that the operation is at peak efficiency. Such data belongs to two categories, surface quality parameters and operational performance parameters each contributing its major share to keeping the standard of production high.

Among the main characteristics of Quality and Performance Data are surface finish quality, defined as the texture and smoothness of the roll's surface after it has undergone grinding. This normally finds expression in quantification by the measurement of surface roughness after grinding. The rolls that go into high-precision applications have to be of a surface finish that is consistent and within tight tolerances since this directly influences downstream processes. Surface roughness is measured in micrometers(μm), and any deviation from the required roughness will impact the performance of the roll. Continuous tracking of surface finish quality ensures that each roll meets specifications and minimizes the possibility of rework or rejection later in the production chain.

Another important property is **dimensional tolerance** or how close the actual size of the roll is to some set of limits. Little differences in the diameter, length, or roundness of the roll might result in issues at later phases of production, or even possibly make the product unsalable. It is for this reason that dimensional precision is such a critical measure. The characteristics monitored to tight tolerances are the overall diameter of the roll when it has been ground, and the grinding error, which is represented as the difference between the actual and desired dimensions of the roll. By utilizing these characteristics, the operator will know if the grinding is taking place at the expected level of accuracy. thus, should there be any deviations, they must be corrected immediately.

Besides quality metrics, other performance related data such as cycle time and throughput provide important indications of the productivity of the grinding machine. The term **cycle time** refers to the length of time between setup for one complete grinding operation and final inspection, while throughput deals with the total number of rolls processed within a certain timeframe. Reduction of cycle time to the barest minimum, without sacrificing quality, allows more production within the same timeframe and enhances the efficiency of the grinding operation in general. This is the actual time spent in grinding, a subset of cycle time that provides a granularity in understanding the time efficiency of the process. It is used to analyze the grinding time with other parameters for the same type and can help manufacturers to identify whether bottlenecks or inefficiency in the grinding process exist.

Another relevant measure is **tool wear rate**, which is basically the rate at which deterioration of the grinding tool, in most cases, a grinding wheel occurs [27]. In addition to increasing operating costs through the frequency of the tools' replacement, extreme tool wear results in a poor surface finish and dimensional accuracy. In turn, tool wear monitoring will help operators replace wheels at proper times, hence a grinding process will be stable in addition to excellent maintenance. The total number of strokes, or passes, of the grinding wheel across the roll will provide the tool wear and grinding uniformity information. Increased stroke counts might indicate that tools are used beyond their capacities and degraded surface quality and accuracy.

Productivity metrics are product yield and reject rate and, in this regard, are key measures of effectiveness in the grinding operation. **Product yield** refers to the number of rolls which meets all criteria of quality, whereas the **reject rate** is the percent of rolls that do not meet specifications. Often, high reject rates indicate problems with either grinding accuracy or surface finish and provide a trigger to adjust the process. By monitoring these rates, operators would be able to ensure high levels of yield while immediately investigating causes of rising rejections.

The last parameters are the roll grinding operations, including **total roll removal** and the diameter of the roll after grinding. This is the quantification of the total material removed from the roll during grinding, necessarily needed to reach the final dimensional requirements. Roll diameter after grinding is checked against target dimensions to verify the grinding process has reached required thickness and roundness. The latter is very important in applications because rolls need to show a good fit within equipment specifications, and their slightest deviations make the processes further downstream inefficient.

Data collection and analysis of Quality and Performance in Real-time will give a holistic view of the grinding process to maintain high standards of product quality and machine productivity [24]. By continuously tracking surface finish, dimensional accuracy, tool wear, and production efficiency, manufacturers may drive evidence-based decisions to optimize process parameters, minimize rejects, and maximize product yield. This multidimensional approach towards quality and performance monitoring lies at the heart of ensuring a ring of consistency in the quality of rolls produced, reduces waste, and generally economizes the operations of roll grinding.

2.4.3 Maintenance and Condition Monitoring Data

Maintenance and Condition Monitoring Data are the two most significant constituents in the longevity and reliability of the roll grinding machine, thus keeping it fit and updated for predictive and preventive strategies [28]. A dataset of this type will enable manufacturers to trace early signs of wear or malfunction that can be intercepted in time, thereby minimizing unplanned downtime and enhancing the useful life of the machine. By continuously measuring all types of machine health indicators, the maintenance teams could be more proactive in their repairs instead of being bound to their fixed cycles of repair.

Condition monitoring among other aspects, incorporates one of the most important processes associated with vibration analysis [29]. The **measurement of vibration** within the machine helps in the detection of imbalance, misalignment, or wear inside its components. Increased levels of vibration typically serve as an indication for maintenance need, due to the reason that continuous vibration causes acceleration in the wear on the components and reduces general performance. Another condition monitoring technique involving acoustic emission monitoring is somewhat complementary to vibration analyses. This method detects sound frequencies created during the operation of the machines. Specific changes in sound frequencies can indicate grinding irregularities or the onset of wear in critical parts, like bearings or spindles. This early detection can prevent further damage that is not immediately visible in other data types and thus enables early intervention and can help lower potential repair costs.

Temperature monitoring is another critical aspect in condition monitoring [30]. This may be achieved by mounting sensors on important elements of a machine such as motors, spindles, and bearings, to enable the operator to constantly assess the thermal conditions. Normally, temperatures higher than usual or temperatures that fluctuate indicate problems like insufficient lubrication, excessive mechanical load, or the tendency to overheat. Such problems require immediate attention to avoid damage. For instance, high temperatures for a long period in the spindle area could indicate the wear of bearings or some breakdown in the cooling system that would hamper machine performance if left unconsidered.

Vibration, acoustic, and temperature data are complemented by lubrication and oil analysis to further investigate the health of machines [30]. Regular checking of lubrication and analysis of the quality of the oil used in bearings, gears, and hydraulic systems provide early detection of contamination or breakdowns in lubrication. For example, the presence of metal particles or other contaminants in oil samples may give indications of wear and tear within internal parts. This could serve as an early warning that some sort of maintenance is required. This will be predictive maintenance that avoids sudden failure of components for smooth and efficient running of the grinding machine.

Last but not least, **machine load** and **power consumption data** provide fundamental insight into the efficiency of the machine [31]. This helps in identifying problems of a mechanical or operational nature through observing the load on the machine and its power consumption, which are usually invisible at first sight. If the machine begins to use more power than it usually consumes, then that would indicate a potential hidden mechanical problem or the need for adjustments for optimal performance. Tracking the load and power consumption helps in optimizing energy use, thus making the process of grinding more cost-effective.

Maintenance and Condition Monitoring Data are basic in diminishing machine downtime and give a reason for the data-driven maintenance strategy [29]. As long as early warning signs such as increased vibrations, unusual acoustic emissions, elevated temperature, oil contamination, or abnormal power usage are recorded, the manufacturer becomes able to move away from reactive and towards proactive maintenance. With this approach, the equipment uptime and reliability are not only maximized but the maintenance cost is drastically reduced to ensure that the grinding machine remains up, running, and efficient in producing quality rolls.

2.5 Key Challenges in Data Utilization in Roll Grinding

Data utilization plays an important role in process optimization, quality assurance, and predictive maintenance in roll grinding operations. However, the effective usage of data within such a niche also comes with many challenges. Major challenges of data utilization include complexity in collecting data, data integration across systems, handling high-frequency sensor data. In other words, each of these challenges needs to be met if the full power of data-driven strategies in roll grinding is to be unleashed to grant manufacturers the required productivity, quality, and cost efficiency.

2.5.1 Data Collection Complexity

The complication associated with data collection is one of the major problems in data utilization in roll grinding [23]. A number of sensors and control systems are utilized in the different types of roll grinding machines to collect huge levels of operation data, including temperature, vibration, acoustic emissions, force, tool wear, and machine parameters such as feed rate and wheel speed. The diversity of data sources necessitates the use of specialized sensors that can measure the required variable with precision, together with an advanced system of data acquisition. Not every machine is fitted with standardized sensors, and variations in equipment and sensor technologies make certain data collections inconsistent [30]. Besides, environmental conditions in the form of dust, temperature fluctuation, and mechanical noise interfere with the accuracy of the sensor. The periods required for its calibration and maintenance are also a blot on its capability for high-quality data capture.

Yet another layer of complexity in data collection is to determine which data to capture relevant to various stages in a grinding process. It requires some basic understanding of the grinding process itself to determine which variables most influence product quality and machine health, which again may be reached through trials and errors. Besides, grinding is a highly dynamic process. minute-to-minute variations have to be reflected in the data collected, which demands a highly sophisticated data acquisition system that runs at high speed, with similarly sophisticated solutions for data storage. Overall, data gathering in the roll grinding process is complex and needs an integrated approach balancing sensor placement, data precision, and operational feasibility.

2.5.2 Data Integration Across Systems

One major challenge in using data for roll grinding arises when perfect data integration from various systems that deal with production and maintenance is performed [32]. The process of grinding rolls involves numerous different platforms or devices (in general, programmable logic controllers-PLCs, numerical control-NC systems, industrial PCs, and external databases like the History Database) used for long-term data storage. Each can be set up with its protocols, data formats, and even communication standards. This makes integrating the data from all different sources into one full framework quite difficult.

This implies that for the effective use of data being generated, there must be immediate access and inter-operability of data from sensors, control systems, and historical databases across all relevant software and hardware platforms. A variety of integration barriers abound where there are legacy equipment and proprietary software common in manufacturing. For example, the industrial PC hosting the control system might speak a different language from what the cloud-based analytics platform or predictive maintenance software can understand. Thus, real-time aggregation and analysis may be extremely hard. Overcoming such integration barriers often involves dealing with specialized middleware or custom-built data connectors that translate and synchronize data across disparate systems, which is often highly cost and technically complicated.

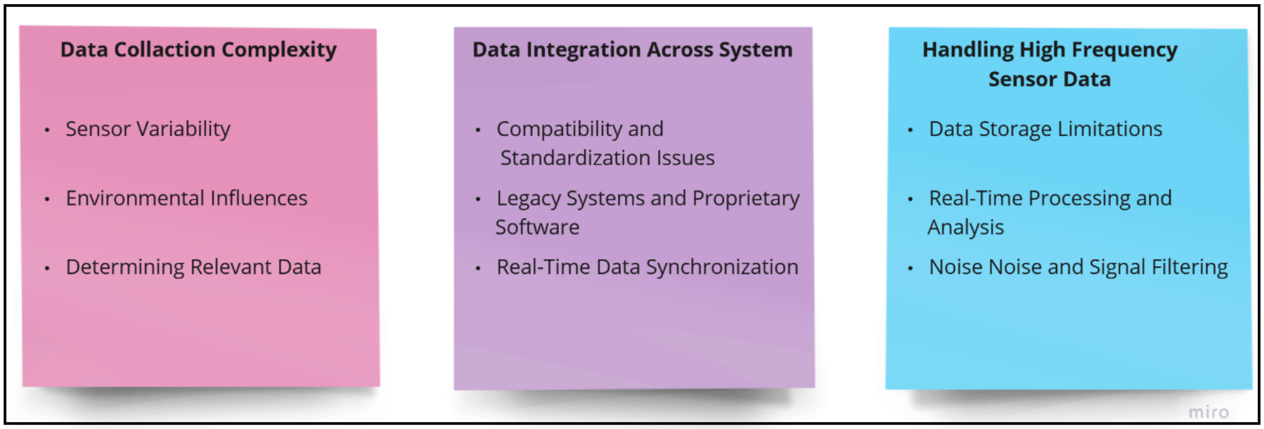

Figure 6: Key Challenges in Data Use in Roll Grinding

Besides, for real-time monitoring and predictive analytics, different systems data needs to be integrated where various data streams have to be processed in a single stream to show insight into the action [19]. It may turn out pretty difficult to monitor the holistic performance of the machines without integrated data systems, delays, and inconsistency in the flow of data leads to missed opportunities regarding the optimization of processes and maintenance interventions. In this respect, effective data integration has a critical role to play in continuous improvements and support for the data-driven approach to roll grinding.

2.5.3 Handling High-Frequency Sensor Data

Among other challenges, there is the important question of how to handle and process high-frequency sensor data emanating from the operation of this machine [32]. These may include high-frequency data such as vibration and acoustic emission signals measured in very short intervals in order to catch minute changes in machine behavior that might indicate wear, misalignment, or process anomalies. While this type of telemetry data is very useful to monitor machine health and product quality, volume and speed create an emergent challenge in data storage, processing, and analysis.

With high-frequency data, robust data storage is required, which can perform on higher scales without lagging the performance of the system [28]. The traditional data storage system may lag behind the demanding nature of high-frequency data and may cause bottlenecks or loss of information in data. Secondly, high-frequency data needs analysis and processing in real time for immediate decisions, which, in turn, again requires superior data analytics tools and immense computational powers. Real-time response capabilities demand immediacy. therefore, when the monitor detects spiking vibrations, any delay in the process may lead to possible damage to the roll or machine components.

High-frequency data is also extremely noisy and requires intensive filtering and processing in order to tease out relevant information [32]. It requires data scientists and engineers to develop algorithms that will uncover relevant patterns or anomalies from the stream of raw sensor data continuously being generated. In fact, managing such high-frequency sensor data does indeed require high-performance computing infrastructure, effective data management strategies, and specialized analytics in order to deliver timely and accurate insights.

2.6 Production Analysis in Roll Grinding Industry

The analysis of production in the roll grinding companies has undergone a radical transformation, moving from the most basic manual inspections and recordation to the most sophisticated data-driven method [25]. This change reflects the paths the industry has taken in response to the need for more accuracy, efficiency, and productivity as well as new developments in machine learning, sensor technology, and real-time data processing. In-depth examinations of the grinding process, machine performance, quality assurance, and predictive maintenance are now included in production analysis.

2.6.1 Historical Overview of Production Analysis in Roll Grinding

Historically, production analysis in the field of roll grinding depended upon manual measurement and observational techniques [33]. In the early years of industrial manufacturing, functional requirements dominated, with very little importance given to precision or repeatability within successive production cycles. Operators conducted basic parameter measurements such as roll diameter, surface finish, and concentricity using micrometers and surface roughness gauges. These measurements, often recorded by hand, typically resulted in poor data quality and gave hardly any information regarding machine performance.

Yet, as manufacturing technology evolved, essential automated control systems were implemented in roll grinding, providing better measurement repeatability and accuracy. Generally, these systems were stand-alone and not integrated to other plant systems or historical databases. Production analysis was commonly realized during this era in a reactive way, that is, problems were corrected once they were noticed, and little was done to predict an issue before it happened and/or avoid it. For example, maintenance was typically scheduled on a fixed interval basis, resulting in either over maintenance or unexpected breakdowns that impinged on overall productivity.

The rise of the CNC systems marked a major turning point in the industry for roll grinding [33]. With the CNC technology, machine parameters could be set to much higher levels of preciseness, and the data from these machines could now be stored and utilized by operators to track performance trends over time. This also laid the groundwork for a more sophisticated production analysis to take place because data coming from these CNC systems could be reviewed and analyzed against process improvements in hopes of reducing waste while improving quality consistency. Even with the advantages of CNC technology, however, the ability to analyze production was then limited by a lack of intercommunication between the various systems, preventing a full, real-time view of the entire grinding process.

2.6.2 Modern Analytical Approaches in Roll Grinding Machines

The last couple of decades have seen the rollout of a sea change in the roll grinding industry with state-of-the-art analytical methods enabled by modern sensor technology, data analytics, and machine learning techniques. Modern roll grinding machines come fitted with a host of sensors which record continuously high-frequency data on machine parameters, environmental conditions, and product quality of the finished rolls. This data is treated in real time and stored in integrated databases, thus constituting a strong basis for continuous production analytics.

The modern approaches used for analysis in the grinding of rolls can be broadly divided into different categories:

Real-Time Monitoring [23]**:** As with any production analysis, most modern roll grinding machines have employed real-time monitoring of performance. Machines fitted with IoT-enabled sensors are in continuous capture of critical parameter data, such as vibration, temperature, force, and spindle speed, etc. This real-time data enables operators to see instant machine performance and serves as an early warning for drifts or deviations from optimal set conditions. For instance, if the vibration suddenly increases, that would point to imbalance or misalignment, and thus operators may make an adjustment before it could impact grinding quality or cause damage to machine components. Real-time monitoring also provides immediate checks on quality, in terms of parameters related to surface roughness and dimensional accuracy, being continuously reviewed against tolerance requirements.

Digital Twins and Simulation: One of the fast-emerging approaches in production analysis is the use of so-called digital twins, virtual replicas of physical roll grinding machines. In this respect, the digital twin simulates a number of machine conditions in which operators can try various grinding scenarios, processes, and maintenance, all without disturbing real production. Data from the physical machine is constantly feeding into the digital twin model, updating it in real time. It enables production planning, optimization of machine settings, and the behavior of machines under varying conditions for further preventive maintenance and process improvements.

Integration with Advanced Analytics Platforms [23]**:** Most modern roll grinding machines are increasingly integrated with advanced analytics platforms that aggregate data from multiple sources for in-depth analysis. The platforms provide the capability for data visualization, automated reporting, and cross-functional insights that surpass simple machine metrics to broader production and quality performance indicators. It may, however, be an analytics platform that will bring out even trends across shifts or batches and serve to improve resource allocation, load balancing, and inventory management. Through such integrated systems, a more holistic approach can be provided for production analysis, using the data to drive on improvements through the whole manufacturing line.

2.7 Applications of Predictive Analytics in Roll Grinding

Predictive analytics in this specialized field of roll grinding allow for the derivation of critical insight into machine performance optimization, product quality improvement, and proactive maintenance strategies [23]. Probably the best manifestation can be found in predictive maintenance, where these data-acquiring sensors can give an early warning in terms of wear and tear, misalignment, or the possibility of failure through the three common types of inputs vibration, temperature, and acoustic emission. These are the data points analyzed to enable predictive models that allow detection of maintenance needs in advance, prior to any actual breakdown, enabling the operators to schedule repairs at an optimum time. The concept here ensures that there is minimal unplanned downtime, such a policy prolongs equipment operational lifetime by preventing heavy damage to critical components.

Another critical application of Predictive Analytics relates to quality control and consistency. Precise surface finish and dimensional accuracy are paramount in roll grinding, as subsequent processes depend directly on the same. Predictive models help in monitoring real-time variations in grinding parameters. If any deviation in wheel speed, feed rate, or depth of cut is observed, then immediate remedial action may be undertaken to maintain the quality of every roll. This real-time oversight greatly reduces errors and rework, which contributes to consistent, high-quality production results.

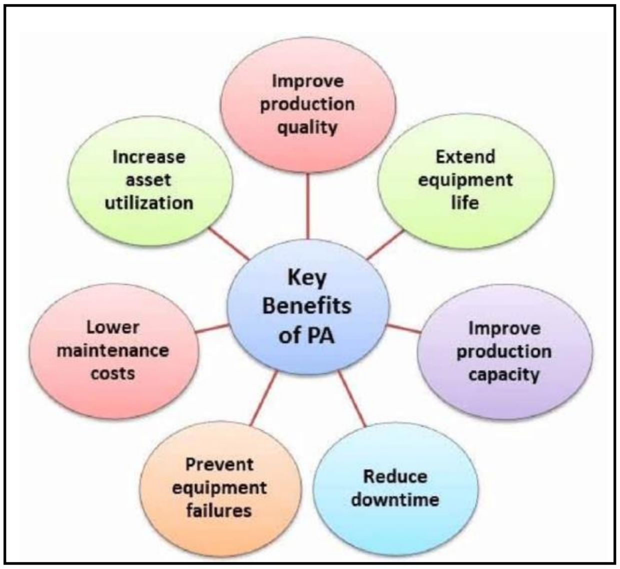

Figure 7: Key Applications of Predictive Analytics [35]

Another area where predictive analytics demonstrates real value is in the optimization of grinding parameters [34]. If historical grinding cycles were monitored and recorded, a predictive model can identify, from prior experience with a given material and roll specifications, the optimal wheel speed, feed rate, and coolant flow rate. By constantly manipulating such parameters, which change with the conditions, the operators maintain a stable and efficient process, minimizing variability, reducing cycle time, and optimizing the usage of resources. The result is that the grinding operation becomes more predictable and under control result important for meeting production quotas without sacrificing quality.

Predictive analytics also strongly enhances tool wear management [27]. Analyzing the wear patterns of grinding wheels and other tooling components, predictive models can estimate when replacements will be required to help operators plan and avoid unexpected tool failures that may compromise quality or halt production. This proactive approach with tool wear minimizes downtime from unplanned changes and supports cost control by extending the usable life of consumable tools. Accordingly, operators can ensure grinding quality and continuity of operations.

Additionally, predictive analytics for energy efficiency and reduction in grinding time determine the best settings for time and power consumption [31]. Predictive models based on historical data are able to define what configuration will reduce the grinding times without affecting the quality of the product and the adjustments that result in minimum power consumption. This becomes increasingly important in today's manufacturing environment, where energy efficiency is both an environmental and cost consideration. By streamlining grinding times and reducing energy consumption, predictive analytics support sustainable production goals by reducing operational costs.

2.8 Existing Gaps

While predictive analytics has shown enormous potential in improving the grinding operation, there remain some gaps that prevent its complete implementation and effectiveness. These need to be specifically addressed for furtherance of the industry in process optimization, machine reliability, and consistent quality.

2.8.1 Lack of Comprehensive Predictive Models

The unavailability of specific predictive models based on production parameter data concerning the very industry is a significant restraint toward the application of predictive analytics in the area of roll grinding. No universally applicable model has been developed so far to predict the critical operational parameters of the process of grinding rolls, and such a limitation restricts the effectiveness of predictive maintenance and optimization efforts. Predictive models available normally cover larger manufacturing processes or specific machine types and thus often lack the granularity to address unique dynamics inherent in roll grinding.

The absence of such focused predictive modeling relates to issues such as the complexity and variability of parameters involved in roll grinding, including dimensions of the rolls, grinding force, and machine wear rates [23]. All these factors vary substantially from one process to another and require flexible models capable of considering frequent changes in conditions of production, as well as the properties of materials. Predictive models for this domain often make use of little or isolated data, since they are often based on historic data from just one machine or line of production. This reduces their generalization capability across different machines or industrial settings and thus limits their possibility of getting wider operational insights.

Without an all-inclusive model to interpret roll grinding production parameters, a facility would be limited in what they could do on proactive adjustments with predictive data. Solving this gap requires the development of sophisticated machine learning models that have the capability to process large volumes of data from multiple sources, including near real-time data from sensors, historical performance metrics, and material specifications. These models will have to use adaptability as continuous learning so that they can have the capability to handle their objective with accuracy under various operational conditions.

2.8.2 Data Integration Issues Across Systems

Another important gap in predictive analytics adoption pertains to the integration of data emanating from diverse systems operating in the process [32]. Some of the sources from which data emanates in a modern roll grinding process include sensors, control systems, quality assurance tools, and maintenance logs. These systems, many times stand-alone and do not share more than the minimum amount of information, so that one can easily see how it results in fragmented data silos. In such cases, information provided by the same type of machine may be stored in formats that are different from others or in different databases, and this creates huge problems when it comes to integration. Sometimes, interoperability is hard to achieve since different manufacturers of equipment may use a certain type of proprietary communications protocol that makes standardizing the collection of data quite hard. Besides scalability issues, due to most facilities lacking the infrastructure to efficiently store, process, and analyze increasingly growing volumes of data, many of these facilities also face scalability issues. This further aggravates the problems so that point solutions for predictive analytics are confined to localized areas of the production process rather than offering a holistic view into machine performance and operational health. Overcoming all these challenges will be hinged on the full realization of the power of predictive analytics in the optimization of roll grinding operations.

The above discussion outlines the few important challenges and gaps in the existing landscape on data utilization, predictive analytics, and production optimization in roll grinding operations: from a lack of comprehensive predictive models to integration with disparate systems, and handling high-frequency sensor data. These gaps suggest that there is a strong need for innovative solutions that can translate recent advances in machine learning and data analytics into efficiency and quality improvements in grinding processes. This thesis seeks to address these gaps by developing and implementing data-driven methodologies tailored to the unique requirements of roll grinding. In particular, the approach involves real-time sensor data, advanced predictive models such as Long Short-Term Memory (LSTM) networks and ensemble machine learning models, and sophisticated preprocessing techniques to deal with diverse and high-frequency data sets.

The subsequent chapters will thoroughly explore the entire research process, starting with the theoretical background, the background will address statistical approaches for modeling, offering insights into the statistical tools and methodologies that enhance the accuracy of predictions. Theoretical discussions will also encompass a comprehensive review of evaluation metrics for predictive models, focusing on how performance is assessed and how these metrics guide model selection and enhancement.

The methodology chapter will provide a detailed account of the research process. It will start with an in-depth description of the data sources used in the thesis, including how the data was collected and the criteria for selecting the datasets. The chapter will then cover the data cleaning process, emphasizing the steps taken to address missing values, outliers, and inconsistencies. Following this, the experimental design will be outlined, offering insights into how the experiments were structured to test the research hypotheses. The data transformation techniques used to prepare the data for modeling will also be discussed.

The results section will showcase the outcomes of all models assessed in the thesis. This section will compare the performance of various models, providing quantitative results for each and demonstrating how they perform under different conditions. It will include a thorough analysis of model accuracy, precision, recall, and other relevant performance metrics, facilitating a clear understanding of the findings.

The last sections will present discussions on the real-world implications of the results and how the proposed methods contribute to the optimization of roll grinding operations, bridging the gaps identified in current practices. The discussions will explore how these advancements address existing challenges, improve operational efficiency, and bridge gaps identified in current industry practices. Ultimately, the research aims to offer valuable insights that enhance both the theoretical understanding and practical implementation of roll grinding optimization.

3. Theoretical Background

This chapter lays out the fundamental theoretical background, concentrating on the key elements necessary for creating and assessing predictive models. It starts with an overview of important prediction techniques, showcasing commonly used methods and their applications in addressing real-world challenges. Following that, it explores statistical methods for modeling, focusing on the tools and techniques that facilitate accurate representation and analysis of intricate data patterns. The chapter also covers evaluation metrics for predictive models, offering a framework for measuring model performance and ensuring their practical relevance and reliability in industrial settings. This theoretical foundation provides a solid basis for the following methodology and analysis.

3.1 Key Prediction Techniques

3.1.1 Supervised Learning

Supervised learning is one of the important paradigms of machine learning, whereby systems execute tasks based on labeled datasets [37]. Correspondingly, each input has a correspondence in relation to a target output value. Therefore, one of the main goals is to learn an appropriate mapping function from inputs to outputs so that it generalizes well to unseen data. Several applications implement this approach, including image recognition, fraud detection, and medical diagnosis. Supervised learning, by reducing the prediction error on labeled data, tries to give a high degree of accuracy with high reliability. It essentially comprises mainly two stages: during a training phase when it discovers patterns from the input data, and during a testing phase that assesses its performance on unseen data. Key techniques in the domain of supervised learning include complex tasked specialized algorithms including neural networks, ensemble methods that allow robust predictions, and more lightweight but highly effective models such as decision trees and support vector machines. Each algorithm has a different fit for certain data types and problem areas, and appropriate choices between them lie along the spectrum of multimodal toolkit use to solve large-scope problems.

3.1.1.1 Long Short-Term Memory (LSTM) Networks

LSTMs are a type of RNN that is designed for effective prediction and processing of sequential or time-dependent data [37]. Traditional RNNs cannot avoid the vanishing gradient problem that makes them incapable of capturing long-term dependencies in sequences. This is mitigated in LSTMs by a very special architecture based on memory cells and gating mechanisms that regulate information flow within the network.

Memory cells are utilized to provide a sort of long-term memory for the network by storing information for long periods of time. During training, these cells determine which information gets stored, updated, or discarded according to the specific learned patterns of the data. The input and output of such cells are controlled by gates: the input gate, the forget gate, and the output gate.

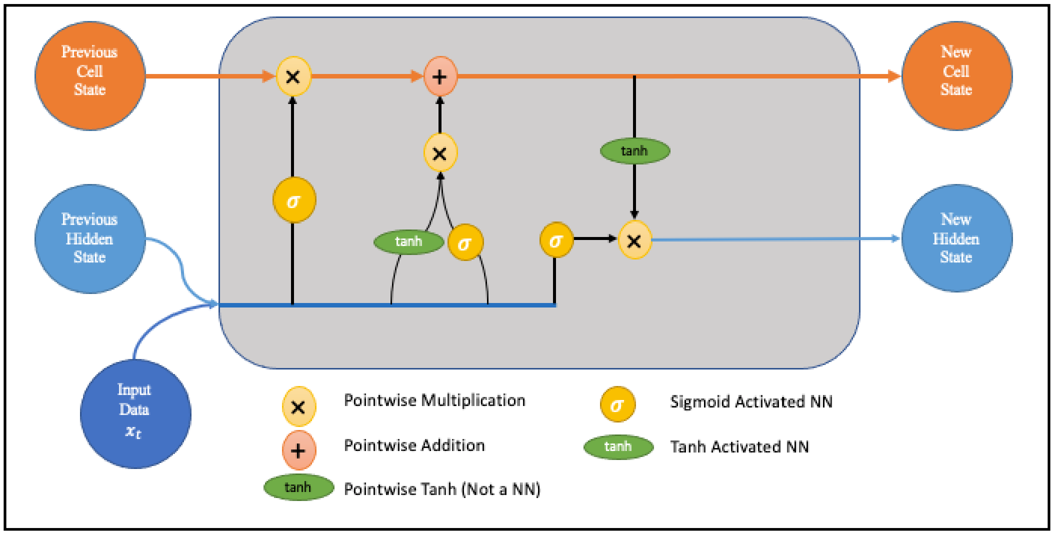

Figure 8: Overview of LSTM Model [37]

The input gate regulates what new information is stored in the cell. It is the forget gate that decides what previously stored information should be erased [39]. The output gate determines what amount of the memory cell's content will affect the current output. This strong mechanism provides the possibility for LSTMs to learn both short and long-term dependencies in sequences, allowing them to perform particularly well in tasks with rich temporal relations. The LSTM model is especially effective for predicting the parameters of roll grinding machines because it can capture temporal dependencies and long-term patterns in sequential data. By processing time-series data efficiently, LSTM can model the dynamic relationships between machine parameters, leading to accurate predictions that enhance grinding performance and minimize operational inefficiencies.

3.1.1.2 Gradient Boosting Regressor (GBR)

The Gradient Boosting Regressor is the most powerful supervised learning algorithm in regression tasks [40]. It works on the boosting principle, where weak learners, usually decision trees, are trained sequentially to minimize errors made by previous models. As showed in figure 9, each new model corrects the residuals of the previous models, thereby improving the overall predictive performance. GBR cascades several weak learners into a strong predictive model that generally performs well in modeling complex relationships between data. It further uses a differentiable loss function, such as MSE, for optimization, hence the model iteratively refines its predictions.

One major strength of GBR is the possibility to optimize multiple hyperparameters containing estimators as in terms of number of trees, the learning rate, depth of tree, and subsampling rate. Such parameters regulate bias-to-variance tradeoff and so enables it to be well performed by the model. Even though GBR was quite successful, the weakness that follows in GBR was that its response to outliers as well as very long times to train than those compared with a more complex model. While handling non-linear relations and high predictability is very important, its general popularity for the applications like financial forecasting, health analytics, or engineering makes this model more common.

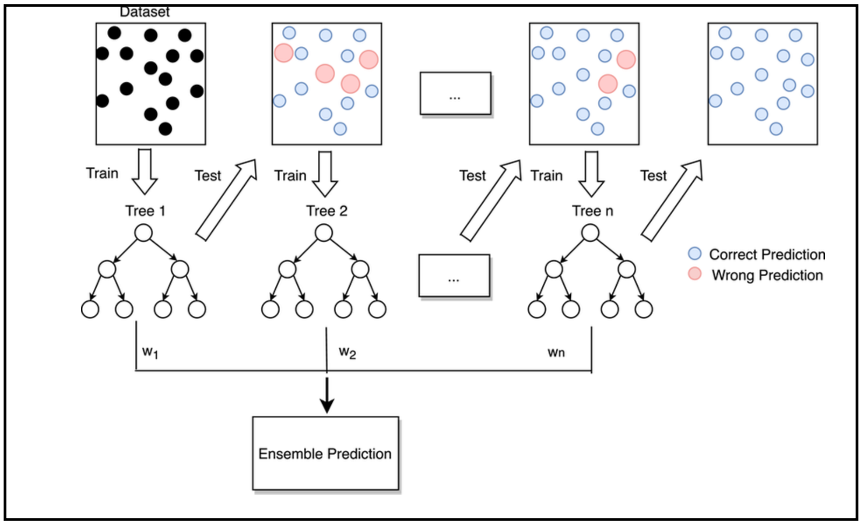

Figure 9: Flow Diagram of Gradient Boosting Regressor [40]

The GBR is particularly well-suited for predicting roll grinding machine parameters due to its ability to handle complex, non-linear relationships in data effectively. The roll grinding machines operate under very variable conditions where various factors such as material properties, operational parameters, and environmental influences interact intricately. GBR is very effective in capturing these interactions by iteratively building models that focus on minimizing the residual errors of previous iterations, thereby uncovering subtle patterns in the data. Its adaptability to different data distributions and robustness in managing multicollinearity make it a powerful tool for modeling the complex dynamics of roll grinding processes.

3.1.1.3 Random Forests (RF)

Random Forests are a kind of ensemble learning, they work by enhancing the predictive performance of decision trees by combining many trees into a "forest." Each tree in the forest is trained on a random subset of the data, a process called bagging, using random subsets of features at each split to introduce diversity and reduce overfitting [42]. The predictions of all trees are aggregated using majority voting in case of classification tasks, and averaging the outputs in case of regression tasks.

Another reason is that random forests are really robust to noise and missing data, since the individual trees can handle partial information [43]. For feature selection, they are very useful since they give an estimation of feature importance that measures how each feature contributes to the model decision making process. It finds huge applications in medical diagnosis for identifying risk factors of diseases, fraud detection by determining unusual patterns, and predictive modeling in structured datasets. In fact, it is one of the most versatile algorithms in machine learning due to its interpretability and resilience to overfitting.

The RF model is particularly effective for predicting the parameters of roll grinding machines because of its robustness and capacity to manage complex, non-linear relationships within the data. By utilizing ensemble learning with multiple decision trees, RF delivers accurate and dependable predictions, making it an excellent choice for optimizing machine performance and improving operational efficiency.

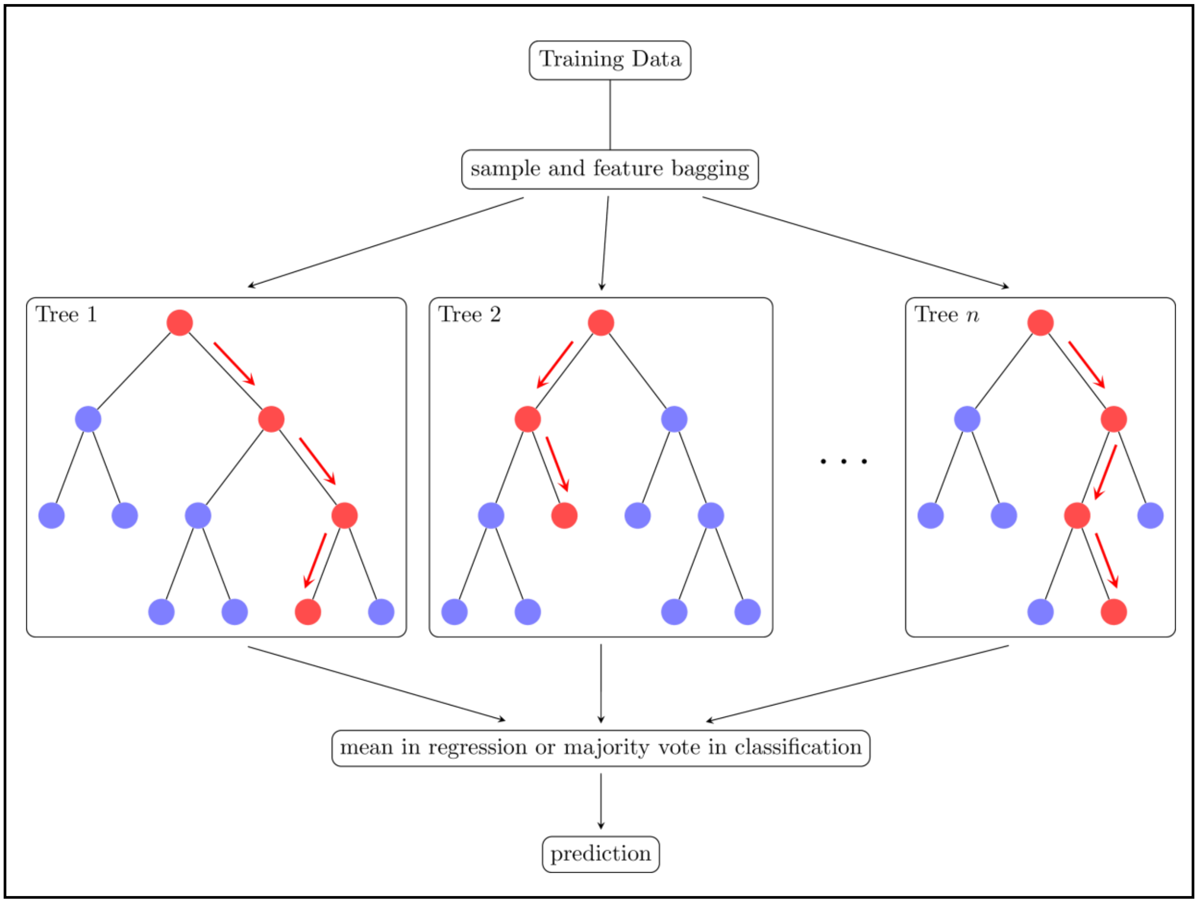

Figure 10: Random Forests Model [42]

3.1.1.4 XGBoost

XGBoost or Extreme Gradient Boosting is a very efficient, high-performance boosting algorithm that builds decision trees one after another in a sequence to minimize the value of an objective function [45]. Unlike bagging-based approaches like Random Forests, XGBoost improves its performance with each subsequent tree in the sequence minimizing residual error from previous trees. It follows advanced techniques such as regularization for penalizing complex models and thereby reducing overfitting, tree pruning to prevent overgrowth of each tree, and parallelized computation for speedier training.

XGBoost has support for sparse matrix handling out of the box and allows for early stopping, a must have when dealing with high dimensional datasets where the model has stopped improving. It has become one of the goes to algorithms in high stakes predictive tasks such as credit scoring, where the accuracy of the risk assessment is key, and also in recommendation systems, where it improves personalization. Its scalability and adaptability also make it a favorite in competitive machine learning and production systems operating on large-scale data sets.

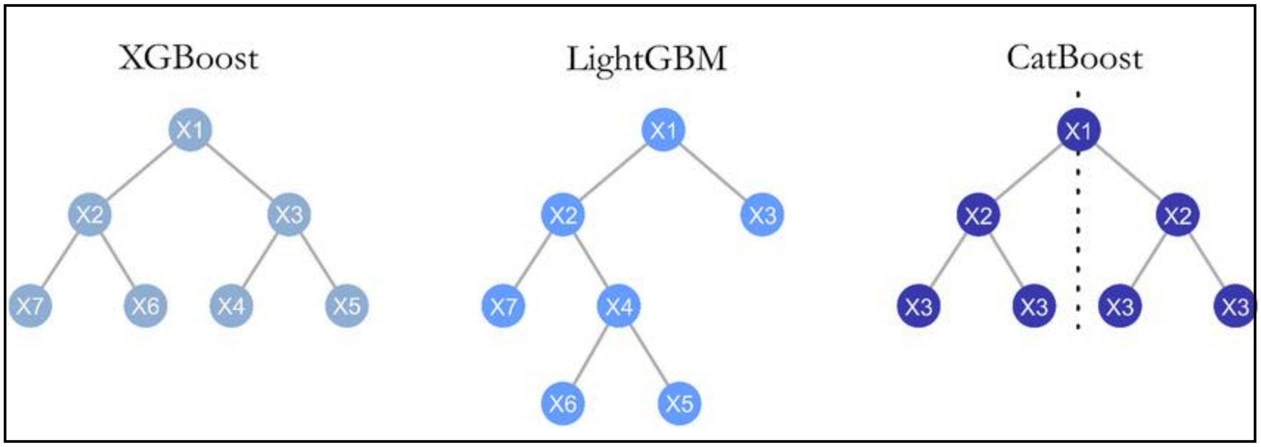

Figure 11: XGBoost model, LightGBM model, CatBoost model [44]

3.1.1.5 LightGBM

LightGBM, Light Gradient Boosting Machine, is an advanced framework of gradient boosting-oriented speed and efficiency, especially on large-scale data [46]. Unlike the traditional algorithms of gradient boosting, in LightGBM, trees grow by a leaf wise strategy, making it actually focus more on the regions of high error and yielding deeper but more accurate trees.

Among the standout features are native support for categorical features that avoid the need for one-hot encoding or any other preprocessing steps of information and save time without losses. It is highly computationally efficient and therefore very well fitted for large-scale applications in finance, such as stock price prediction, in marketing for customer segmentation, and in healthcare for large-scale disease risk analysis. LightGBM allows dealing with millions of points easily, so it is preferred in industry while developing real-time and scalable systems.

3.1.1.6 CatBoost

Yet another gradient boosting algorithm, CatBoost is uniquely designed to handle categorical data very effectively [47]. It uses an innovative technique of encoding for categorical variables at the training time by using combinations of features to capture complex patterns without one-hot encoding or label encoding. A capability that has its advantage in contributing to minimum preprocessing with a reduced risk of overfitting due to poor feature encoding.

CatBoost also has symmetric tree structure and ordered boosting features that will contribute much in training more efficiently and accurately. It is perfectly suited for such applications as customer segmentation, recommendation systems in e-commerce, where categorical data dominates user behavior and product attributes. Moreover, the CatBoost technology outperforms other methods on click through rate prediction for online ads, while its applications are extended to financial analytics for predictive modeling and risk assessment. This level of simplicity and power has made it a go to tool for business applications involving complicated categorical data.

XGBoost, CatBoost, and LightGBM are robust gradient boosting models that excel at predicting roll grinding machine parameters. They are particularly effective because they can efficiently manage complex, non-linear relationships in data.

3.1.1.7 Support Vector Machines (SVM)

Support Vector Machines are a type of supervising learning algorithm with the goal of finding the best hyperplane that separates classes in feature space to maximum margin between classes [50]. This is undoubtedly one of the main reasons why SVM is so powerful, especially when the classes in a dataset are well separated. For problems with nonlinear boundaries, SVM uses such kernel functions as polynomial, radial basis function, and sigmoid to map the data into higher dimensions so that it supports the capability to handle complex patterns.

The SVM remains powerfully immune to high-dimensional data and gives good results even on small datasets, which makes this algorithm a great fit for bioinformatics problems since the datasets contain a huge number of features with relatively few samples. It is also used in image recognition, spam detection, and text classification tasks [50]. In real world practice, however, sometimes too large datasets have to face high computational complexity standing in the way, while its precision and ability to work well in specialized domains make it very useful in machine learning.

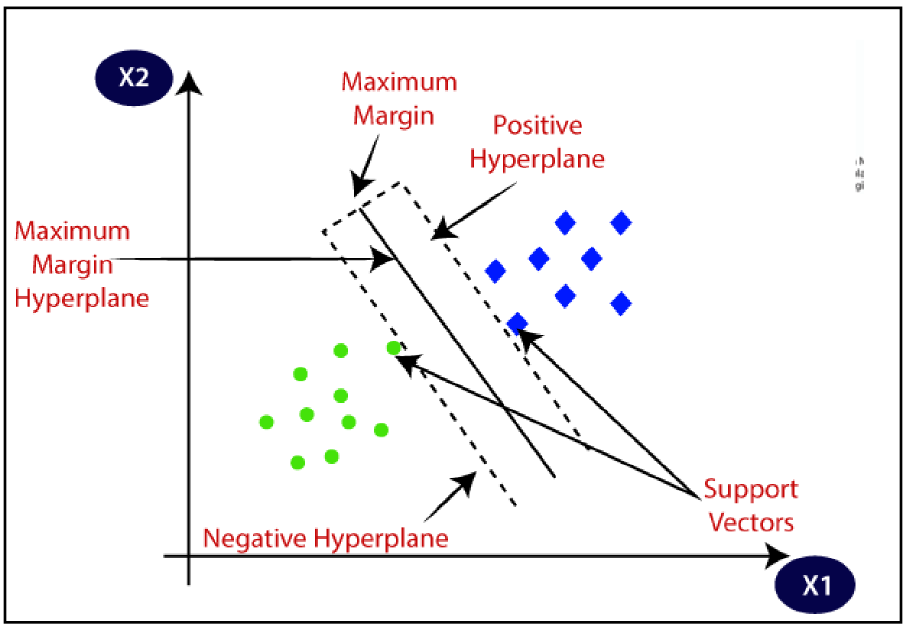

Figure 12: Linear Support Vector Machine [48].

3.1.1.8 K-Nearest Neighbors (KNN)

K-Nearest Neighbors is a straightforward yet efficient instance-based learning algorithm that predicts the class of a new data point by taking the majority class among its k-nearest neighbors in feature space [50]. Unlike most of the machine learning models, KNN does not explicitly undergo any training phase but instead depends on distances, for example Euclidean and Manhattan between data points to make a prediction. The KNN model is useful for predicting the parameters of roll grinding machines because of its straightforwardness and ability to effectively capture local patterns in the data.

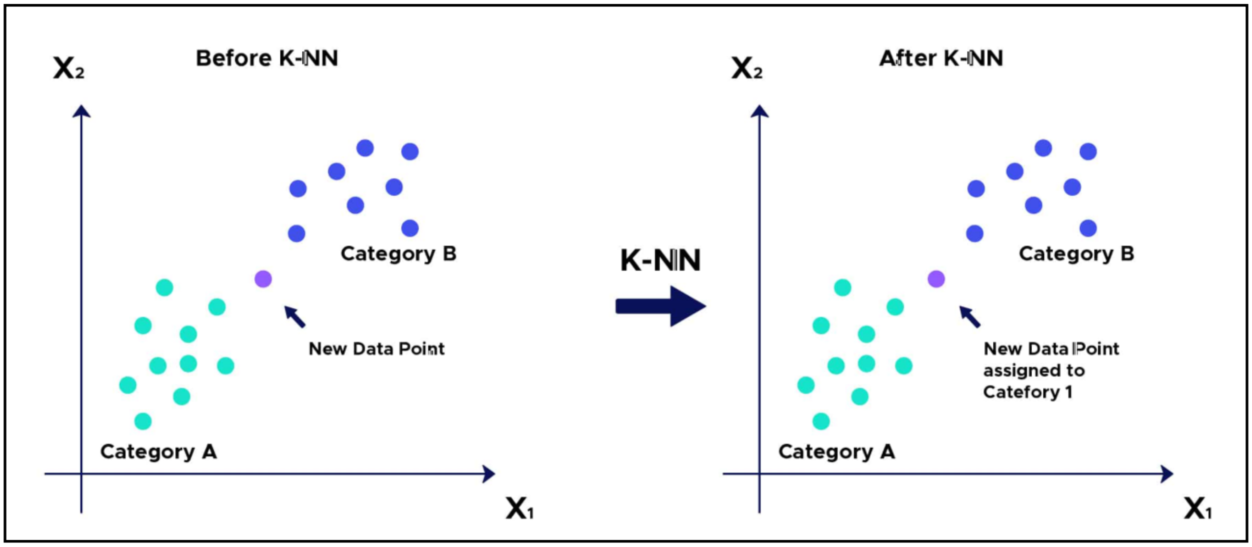

Figure 13: K-Nearest Neighbors [49].

Choice of k and distance metric significantly impact model performance. KNN does well in an application like handwriting recognition, where small scale differences in feature space are to be crucial. And recommendation systems where closeness in feature space may indicate user preference. It finds its usage in anomaly detection for catching outliers. Considering its simplicity, the KNN can be computationally very intensive on large datasets, as predictions require a comparison of the input point to every data point in the dataset [50]. However, its intuitive approach combined with its effectiveness in small scale problems paves the way for its application in many fields.

3.2 Statistical Approaches for Modeling

The statistical approaches to modeling are extremely fundamental in understanding the relationships that exist among variables within datasets, optimizing models, and making predictions based on data. In this section, we look into two of the most important aspects of statistical modeling that greatly impact the performance and interpretability of machine learning algorithms: feature engineering for machine parameters and dimensionality reduction techniques. Both of these methods deal with problems related to high-dimension data, complicated variable relationships, and the need for fast computation.

3.2.1 Feature Engineering for Machine Parameters

Feature engineering is the process of transforming raw data into more appropriate forms required by a machine learning model. In other words, with respect to statistical modeling, it may be defined as the process of new features extraction from the raw data set to enhance the capability of a model dealing with various problems such as multicollinearity, noise, and overfitting issues, and also to grasp underlying patterns and raise predictive performance.

Feature engineering usually consists of several phases of feature selection, transformation, and scaling. Problem and data type and domain expertise will frequently be in control during these phases.

Feature Selection [51]**:**
Feature selection refers to the selection of the most informative features about the target variable by eliminating features that have redundant information or that are not useful. This reduces the risk of overfitting, dimensions, and computational costs. Both Sequential Forward Selection and Mutual Information are variants of a broad category.

Sequential Forward Selection (SFS): Sequential Forward Selection (SFS) is a widely used greedy, iterative approach to feature selection. It begins with an empty set of features and iteratively adds one at a time, focusing on the highest model performance at each step. In every step, the algorithm examines all features not yet introduced into the set and chooses that feature which, when added to the existing set, results in the largest improvement according to some predefined performance metric, such as accuracy or R-squared. It does so until a stopping criterion the inability to improve performance further or until a predefined number of features have been selected-is reached. SFS is particularly effective for small to medium-sized datasets, thanks to its simplicity in implementation, focusing on the most relevant features, which in turn directly enhances the performance of the model.

However, SFS itself has limitations. It can be computationally expensive for large datasets, as at each iteration, it evaluates the addition of every remaining feature. Additionally, SFS only works in a forward manner, thus having a local optima issue. That is, once a feature has been selected, it is never reconsidered in light of later selections that may prove to make for a less-than-ideal combination of features. Despite these limitations, SFS continues to be among the more popular methods owing to its simplicity and how it works rather well to enhance the results of many models.

Mutual Information (MI): Mutual Information is the statistical measure used in feature selection that describes the dependency between one feature and the target variable. It evaluates the quantity of information a given feature would provide about the target and assigns a score to each feature based on this very dependency. Features with higher MI scores are more informative and thus considered more relevant for the modeling task. Unlike measures of correlation that can only point out linear relationships, MI can detect both linear and nonlinear associations, thus making it a very versatile method for feature selection.

The strengths of MI lie in the fact that it is simple and effective, especially in high-dimensional datasets where it serves as a filter method to rank features before more complex feature selection or modeling techniques are applied. Yet, MI does not consider the interactions between features; hence, in some cases when the feature combinations may be important, it yields suboptimal feature sets. Despite this fact, MI is especially good as a preliminary step in feature selection; this helps to reduce the number of features selected on the basis of the most informative one for predictions of a target variable.

Feature Scaling:
The feature scaling in machine learning is a preprocessing step that makes sure all the features have an equal contribution to the model [52]. Most machine learning algorithms are sensitive to the

magnitude of features, and unscaled data can result in biased outcomes or slow convergence during training. Feature scaling usually standardizes or normalizes the data such that all features come to a similar scale. Two common techniques include 'Standardization' and 'Min-Max Scaling'.

Standardization changes features to the mean zero and a standard deviation of one while keeping the distribution of the data, centering it. This is effective in many algorithms assuming a Gaussian distribution, such as Support Vector Machines or Principal Component Analysis. Min-Max Scaling, on the other hand, rescales features to a fixed range, usually from [0, 1]. This helps in algorithms like gradient descent, which converge faster in a bounded range of values.

It enhances the performance of the model through improvement of numerical stability and prohibits features of larger magnitudes from dominating the process of learning. Scaling does not affect the relationship between features, and proper precautions must be observed so that test data are scaled in the same way as training data. Feature scaling is simple but indispensable in many machine learning workflows.

Feature Transformation:
Feature transformation: transformation applied to convert data into a more usable form by increasing linearity, reducing skewness, or features interaction that can influence modeling tasks [56]. Typical examples of feature transformation include logarithmic transformations, polynomial features, power transformations.

Logarithmic transformations normalize heavy-skewed data or variables with exponential growth into a more feasible form for methods assuming normality. Polynomial transformations create new features by introducing interactions or non-linear relationships between variables, increasing their predictive powers. The power transformations, such as those proposed by either the Box-Cox or Yeo-Johnson methods, variances stabilize the data and more closely approximate that of a normal distribution.

Feature transformation is particularly valuable for improving model interpretability and adapting data to meet algorithmic assumptions. However, transformations may complicate feature relationships and require careful consideration to avoid overfitting. When applied thoughtfully, feature transformation can significantly enhance the effectiveness of machine learning models.

3.2.2 Dimensionality Reduction Techniques

Dimensionality reduction techniques are majorly used during the pre-processing of data when dealing with high-dimensional datasets [53]. It seeks to reduce the number of features while retaining most of the critical information, making the representation of the data much easier, hence increasing efficiency for any machine learning algorithm employed. It resolves some of the significant issues in data analysis, such as the curse of dimensionality, overfitting, and computational inefficiency. It not only makes data analysis easier but also improves model performances and interpretability. The major ones employed include Principal Component Analysis.

Principal Component Analysis [53]**:** Principal Component Analysis is a commonly used linear dimensionality reduction technique that seeks to change high-dimensional data into a smaller set of uncorrelated variables, known as principal components. The components are linear combinations of the original features and are ordered in terms of the variance they capture in the data. PCA is effective mainly in the reduction of redundancy within datasets, discovering hidden patterns, and simplifying complex data representation. Thus, it is a must-have in data pre-processing and analysis.

The PCA process follows a number of sequential steps. Among them is the need to standardize data through normalization so that all features have a zero mean and variance of one since PCA is sensitive to the scale of data. Later, the covariance matrix is calculated to show the variable relationships, how much two features vary together. The process of eigenvalue decomposition is done on the covariance matrix to extract eigenvalues and eigenvectors, where the eigenvectors represent the principal components, and the eigenvalues indicate the variance each component explains. Components are ranked by their respective eigenvalues in a descending manner, and the first component covers the maximum variance within the data. The data is then projected onto the top-ranked principal components so as to get a lower-dimensional representation preserving most of the information.

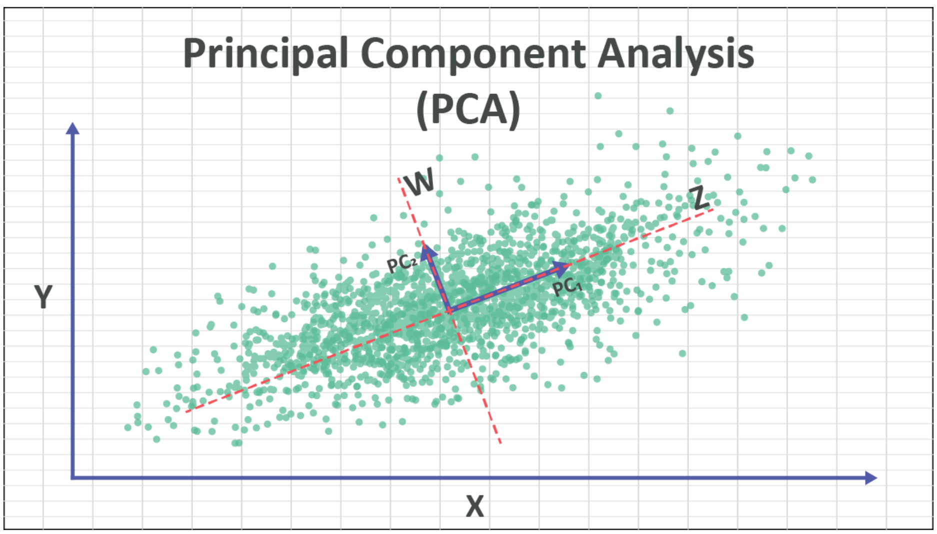

Figure 14: Principal Component Analysis [60]

PCA has quite a number of desirable properties that make it very useful in a wide range of settings. First, PCA diminishes the complexity of high-dimensional data into a more analytically and visually manageable form. Additionally, it reduces noise by concentrating only on those components which explain the most variations, hence filtering out less important variations. Furthermore, PCA improves efficiency by lowering the number of dimensions, thereby accelerating machine learning model training and reducing memory consumption. PCA also helps to remedy a potential problem known as multicollinearity, where redundant information in correlated features decreases model performance by combining correlated features into uncorrelated principal components.

On the other hand, some of the weaknesses associated with PCA are that it assumes linear relationships among variables, which may reduce its performance in case of datasets showing high non-linearities. The PCA consists of the principal components, which are, in fact, combinations of original features and thus often hard to interpret for practical applications [53]. Another problem is to decide on the number of components to retain, retaining too few components may result in the loss of important information, while retaining too many reduces most of the advantages connected with dimensionality reduction.

Applications of PCA cover a very broad range of fields. PCA is utilized in facial recognition and image compression, both of which involve the reduction of dimensionality in image data to retain critical features. The method has also been used in bioinformatics for the analysis of genomic data to find interesting patterns in gene expressions. It is used in marketing and business analytics while customer segmentation and deriving insights on buying behavior. Additionally, PCA is a powerful tool for exploratory data analysis, allowing for the visualization of high-dimensional data in two or three dimensions. PCA is a robust and versatile technique for dimensionality reduction. By capturing the most significant variance in the data, it simplifies complex datasets, enhances computational efficiency, and uncovers underlying structures, making it a cornerstone of modern data science and machine learning workflows.

3.3 Evaluation Metrics for Predictive Models

The evaluation metrics play a critical role in the performance assessment of predictive models. They offer quantitative means to determine the quality of the performance of a model on unseen data and drive how one selects, optimizes, and compares models. The choice of evaluation metric depends on whether a prediction task is done through classification, regression, or other specialized contexts. This section covers some of the main metrics used to assess predictive models. These range from accuracy and precision, mean absolute error (MAE), to the F1 score.

3.3.1 Accuracy and Precision in Predictive Models

One of the simple and widely used metrics for evaluating any classification model is accuracy. It is simply defined as the ratio of correctly predicted observations to the total number of observations. Mathematically, accuracy can be represented as [54]:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

where:

- **TP** (True Positives) is correctly predicted positive observations.
- **TN** (True Negatives) is correctly predicted negative observations.
- **FP** (False Positives) is instances where the model incorrectly predicted the positive class.
- **FN** (False Negatives) is instances where the model incorrectly predicted the negative class.

While accuracy is a direct measure of a model's performance, it can be highly misleading when dealing with imbalanced datasets. In the case of a binary classification problem where 90% of the data points fall into one class, a model that simply predicts the majority class for all samples to attain 90% accuracy gives almost no value from a practical perspective since it completely ignores the minority class.

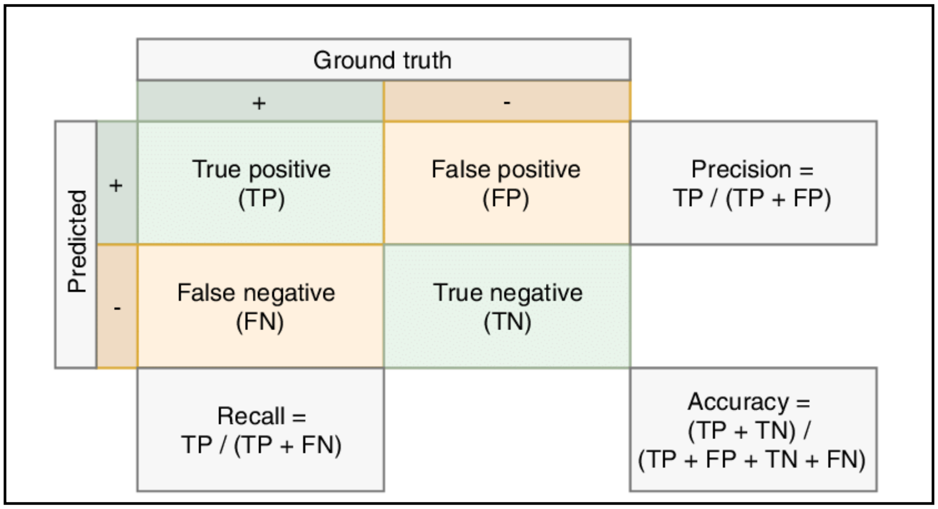

Figure 15: Confusion Matrix [54]

Another important metric, especially in tasks where the cost of false positives is high, is precision. Precision tells us what proportion of the positive observations that were predicted as such were correctly predicted. It is formally defined as:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

Precision, on the other hand, is focused on the quality rather than on the amount of positive predictions. This makes precision really useful in spam detectors, where most legitimate emails could be marked as spam and important emails may never have been read. For instance, in medical diagnosis, high precision implies that positive diagnoses. for example, detecting a disease-have a low level of false alarms, thus reducing unnecessary follow-ups or treatments.

It is complementary to the metric of precision and is often used in conjunction with the recall and F1 score, among others, for providing a more comprehensive understanding with respect to model performance. These two metrics allow a nuanced understanding of how well a classification model can meet the specific demands of an application.

3.3.2 Mean Absolute Error (MAE) for Regression Models

The Mean Absolute Error (MAE) represents a very common and simple metric for evaluating regression models. It relies upon the mean magnitude of errors between predictions and actual values, all in one direction. More formally, MAE is the average of the absolute differences between predictions and ground truth values, which can be expressed as:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i |$$

Where, y_i represents the actual value, $\hat{y}_i$ is the predicted value, and n is the total number of observations.

MAE is straightforward to interpret, as it gives a proper measure of how far the predictions are away from real values on average. One of the most important advantages of MAE involves its resistance against outliers, since it does not square the error values like Mean Squared Error. This also turns out to be one of the serious drawbacks because MAE could be less sensitive to large deviations, and maybe in some cases, that could be a serious minus when large errors are undesirable. In applications where emphasis is on the minimum average error without disproportionately weighting extreme deviations, MAE is particularly useful; hence, it finds a wide application in forecasting and resource allocation.

3.3.3 F1 Score for Classification Models

F1 score is the harmonic mean of precision and recall. It is used widely in classification for tasks involving imbalanced datasets. It provides a global measure that balances the trade-off between precision-or positive predictive value and recall or sensitivity. The F1 score is given by:

$$\mathrm{F1} = 2 \cdot \frac{\mathrm{Precision} \cdot \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}}$$

The F1 score varies from 0 to 1, with values closer to 1 being indicative of better performance. A high F1 score in this case indicates that the model maintains a good balance between precision and recall, both of which are very critical if the applications contain a high cost associated with false positives and false negatives [53]. For example, in medical diagnosis, a high F1 score will ensure that the model minimizes both missed detections (false negatives) and incorrect alarms (false positives).

One of the major advantages of F1 is that it provides more nuance than a strict measure of accuracy, especially when it comes to imbalanced situations. However, the F1 score does not consider true negatives, which can be an important aspect in some tasks [53]. But even with this limitation, the F1 score remains one of the cornerstone metrics for assessing classification models, offering a good way to estimate their performance regarding the prediction of minority classes.

4. Methodology

4.1 Data Sources

Full data collection methods are in place for the manufacturing of machine tools to be done with full efficiency and precision, particularly in the roll grinding process [34]. By implementing data-driven strategies, it will become possible to have continuous monitoring and predictive maintenance that will help improve decision-making processes and enhance product quality. The foregoing sources of data serve as the basic building blocks of the various approaches to advanced manufacturing and provide the dual functionality required for long-term strategic development and short-term operational needs. By applying analytics to uncover key insights, operators and engineers can improve operational effectiveness and the quality of their products.

The information obtained through varied sensors and monitoring mechanisms connected to the machinery is identified and can be placed in different categories. This vast framework allows capturing information on process variability and also helps in achieving preliminary warnings about potential problems to invoke necessary actions. Such records go a long way in attesting to compliance with recognized norms for improvement processes-outputs such as Six Sigma, Lean Manufacturing-requiring broad bases of data-driven analysis towards sustaining improved processes [23].

4.1.1 Overview of Roll Grinding Machine Production Data Sources

The sources of data for roll grinding applications have a vital role in the monitoring, analysis, and optimization of production flows. Large volumes of data are produced from machine tools, which represent all sorts of parameters characterizing performance and conditions during the grinding process. Machine sensors, operational logs, and quality control systems are part of such sources, each with added value regarding insight into the production environment.

Most machine tools are furnished with sophisticated sensor systems that monitor relevant operational variables, such as the spindle speed, feed rate, and position of the tool. With real-time feedback, either the operator or an automatic system can detect deviations from required performance levels. Integrated quality assurance systems also evaluate outputs from grinding operations by measuring dimensional accuracy, surface finish, and other critical parameters.

Most often, data collection in machine tool environments involves a number of systems operating in tandem [19]. While process monitoring systems capture real-time data streams from machine sensors, inspection systems perform the end-product evaluation. All these sources put together form a comprehensive dataset for both operational and quality analyses. Advances in Industry 4.0 technologies, such as IoT and edge computing, have enhanced capabilities for efficient data collection and analyses. These technologies allow seamless integrations between machine tools and their data management systems, hence capture of high-resolution and high-frequency data.

4.1.2 Types of Data Collected

There are various sorts of data that are gathered during roll grinding operations, and each has a specific function in monitoring, analysis, and control. Among these categories are:

a. Operational Parameters

Operational data refers to the dynamic conditions of the grinding process, recording variables like:

• Spindle Speed (RPM): This indicates the rotating speed of the grinding spindle, affecting material removal rate and quality of surface finish. Spindle speed is one of the most critical determinants in grinding efficiency and precision, which affects the interaction of the grinding wheel and workpiece.

• Feed Rate: This is the linear speed of the grinding wheel in relation to the workpiece, which affects the dimensional accuracy and efficiency. The feed rate has to be balanced correctly to avoid excessive wear or damage to the workpiece.

• Head Speed: Denotes the movement of measurement heads or tool systems, critical for real-time process adjustments. Monitoring head speed helps ensure consistent operation and enables the detection of any mechanical anomalies.

b. Dimensional and Geometric Measurements

Dimensional data check the workpiece for its compliance with specified tolerances, including:

• Shape and Roundness Measurements: Measure geometric accuracy to determine the conformance of circularity and flatness. This type of measurement is important to be certain that the product meets engineering specifications.

• Diameter Measurements: Availing critical data on the size of a workpiece at different stages in processing. Consistency in diameter measurements ensures uniformity and proper fit in assembly.

• Eccentricity and Concentricity Metrics: Measure how far gauge alignment can be from ideal about a geometric center. These parameters play an important role in maintaining extremely good accuracies during rotation or fit.

c. Defect Identification and Threshold Monitoring

Advanced machine systems integrate defect detection techniques for surface anomaly detection. Such mechanisms include:

• Crack and Bruise Detection: Sensors detect minor surface flaws in the workpiece and further classify them using a set threshold. To avoid further deterioration and ensure product reliability, cracks and bruising have to be detected at an early production stage.

• Magnetization Tests: Employ magnetic properties as a means of assessing the integrity of the workpiece material. This is especially of importance in materials whose quality and structural integrity could be influenced by magnetic fluctuations in processes.

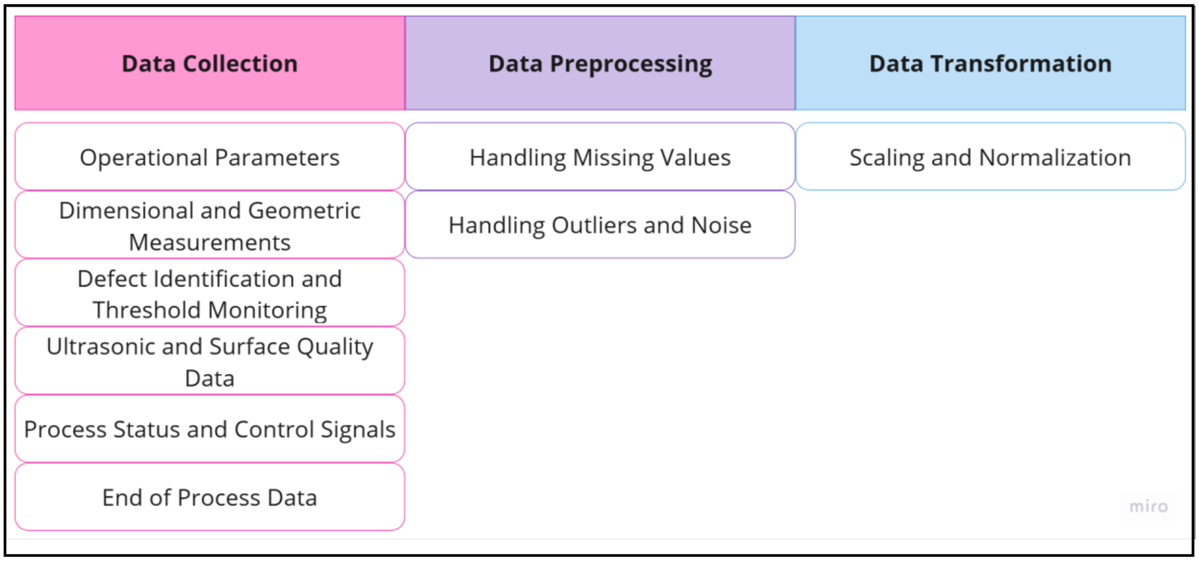

Figure 16: Data Collection and Preprocessing

d. Ultrasonic and Surface Quality Data

Ultrasonic sensors, along with other high-resolution data collection devices, measure surface properties:

• Vertical, radial, and angled amplitude tolerances: The measurement of echo amplitudes in order to assess internal consistency and surface homogeneity. With these measurements, areas are located where possible surface roughness or structural irregularities may lead to the impairment of performance in the product.

• Workpiece Coupling Data: The efficiency of the coupling mechanisms should be evaluated during the ultrasonic test. Proper coupling ensures effective transmission of ultrasonic waves, reducing the chances of incorrect readings and thus providing accurate measurements.

e. Process Status and Control Signals

Operational signals put system states and operator interventions into context. Examples include:

• Pause and Print Commands: These are system- or operator-driven events to indicate a change in the workflow. These signals will be used to control process control and manage production disruption.

• Cycle Completion Flags: Provide an indication of when one processing step is complete and that there is a change in the production cycle. Once cycle completions are detected, operators can manage workflow accordingly, preparing for the next phase of production.

• Tool Remeasurement Indicators: These, depending on wear or process variations, are the causes that have tool systems recalibrated or adjusted. These indicators are important in maintaining accuracy and extending tool life.

f. End-of-Process Data

This section ensures that manufacturing activities run smoothly by recording end-process data such as:

• Final Wheel Measurements: This report is a record of the status of the grinding wheels post-operation. Using this information to assess wear and schedule maintenance ensures that the tool's performance is at its best for subsequent operations.

• Parking Position Indicators: It indicates whether the machine is ready for further maintenance or to be used. Correct positioning of the machine ensures safety and readiness for the next production cycle.

• Features of Process Termination: Categorize the output of the process, so that it is sorted for quality assurance and archival purposes. This information will be useful for future documentation of quality control, audits, and process improvements.

Sophisticated analysis and optimization are only possible with the variety and level of detail of data gathered. In the context of a roll grinding application, it is possible for machine tool operators and engineers to improve overall productivity, reduce waste, and achieve more accurate results using various data types.

4.2 Data Cleaning

Data cleaning is such a critical step in any preprocessing pipeline for data as this ensures that the next analysis has valid and reliable data to fall back on [55]. Obviously, the quality of output for any analytical model would be dependent on its input data. Hence, rectifying discrepancies, missing values, outliers, and noise become crucial. It is important to maintain a high degree of data quality in the context of the roll grinding process due to the complexity of the manufacturing process and the requirement for precise, accurate operational insight. This section describes the techniques used to deal with outliers, noise, and missing values in the dataset.

4.2.1 Handling Missing Values

Effective data cleaning must be performed to ensure that datasets being used for analysis are reflective and accurate for the actual production environment [55]. Among the critical aspects of data preprocessing is how missing values are handled since this can lead to biased analysis with ensuing wrong conclusions.

For missing values, those entries of data entries which contain incomplete records were removed within the provided roll grinding process data. The methodology of imputation with mean or median values is also avoided because of the nature of data collected. These values, involving operational parameters and defect detection metrics in this data set, may show wide variation with changes in machine conditions or stages of the process. Averaging out missing values might mask real variations and may also result in spurious trends not reflecting real operational behavior.

The removal of incomplete entries has ensured that analyses conducted on the data were based solely on complete, unaltered records, thus preserving the integrity of the analysis. This decision was informed by the understanding that accuracy and reliability maintained in the data set is paramount, especially for predictive modeling and quality control analyses where exact measurements are required to guide decisions.

4.2.2 Handling Outliers and Noise

These outliers and noise greatly affect the accuracy and reliability of the data analysis, especially in data related to machine tools, where precision is required [55]. Thus, the identification and mitigation of these are crucial steps toward developing robust models and obtaining meaningful insights from the data.

For handling outliers in the roll grinding data, a thorough analysis was conducted to identify data points that deviated significantly from the expected range. The presence of outliers was assessed through visualization tools such as box plots and scatter plots, which highlighted anomalous values that could distort statistical analyses. These outliers, once spotted, were removed so as not to distort the consistency of the dataset or further the analyses with values that are out of range from normal operating range.

Another point to address was noise in the data, which can be regarded as random variation not bearing any useful information. It can result from many sources: inaccuracies of sensors or temporary fluctuations of the machines' performance. In this project, data points from noise were removed similarly as the outliers. Cleaning noisy data makes sure that only the most relevant and stable data go into further analysis, thus supporting better signal extraction with reduced chances of misinterpretation.

Missing value clearing, outliers, and noise ensured integrity in the analysis of data. This step was, therefore, of utmost importance to get the final datasets for modeling, testing, and reporting as clean and reliable as possible to arrive at more precise and actionable insights into the process of roll grinding.

4.3 Data Transformation

One of the most important stage of the whole preparation pipeline either for machine learning or simply analysis includes a process known as data transformation. In transformation, original data is changed into the most appropriate format or form to be analyzed and possibly transformed into a model. Rather than just optimizing the data towards algorithmic processing, consistency, and interpretability is ensured for more accurate and trustworthy results. Transformation techniques often depend on data type, special demands for analysis, or other conditions of predictive modeling tasks.

4.3.1 Scaling and Normalization

The most important data transformation techniques, scaling and normalization are usually used to change the range of features of data [52]. These processes will ensure that the data conditions are met by several machine learning algorithms, especially those that are sensitive to the magnitude of input values.

For this thesis, standardization is the primary method of scaling and normalization. Standardization is the process of transforming data to fit within a standard normal distribution with a mean of zero and a standard deviation of one. This may be particularly effective in cases where the dataset includes features of different orders of magnitudes. Standardization would rescale the data to such a distribution and minimize any possibility of any one feature dominating the learning algorithm based on its scale.

Mathematically, standardization is expressed as:

$$x_{\text{standardized}} = \frac{x-\mu}{\sigma}$$

where:

- x is an individual data point from the dataset.
- μ is the mean of the dataset.
- σ is the standard deviation of the dataset.

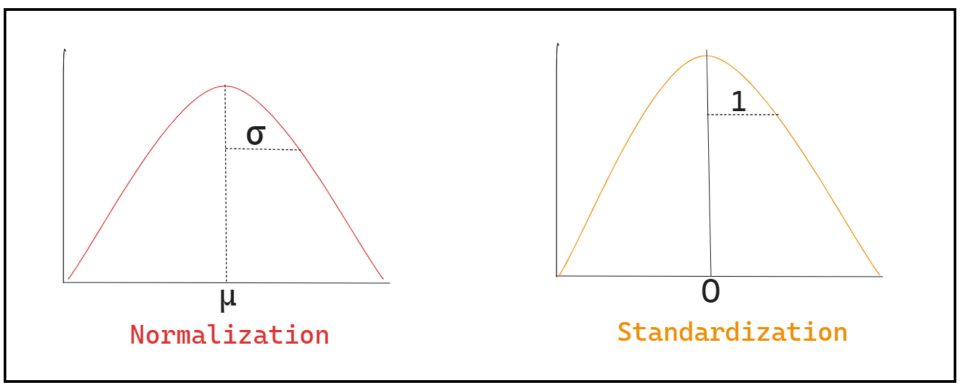

Figure 17: Standardization vs Normalization [60]

However, in this thesis, we also rescaled the data features into a range between [0, 1]. Normalization has the effect of bounding all data points within a specific boundary. This makes them more comparable and helps some models that are sensitive to the scale of features, such as neural networks. It also serves to prevent numerical instability in optimization algorithms when training.

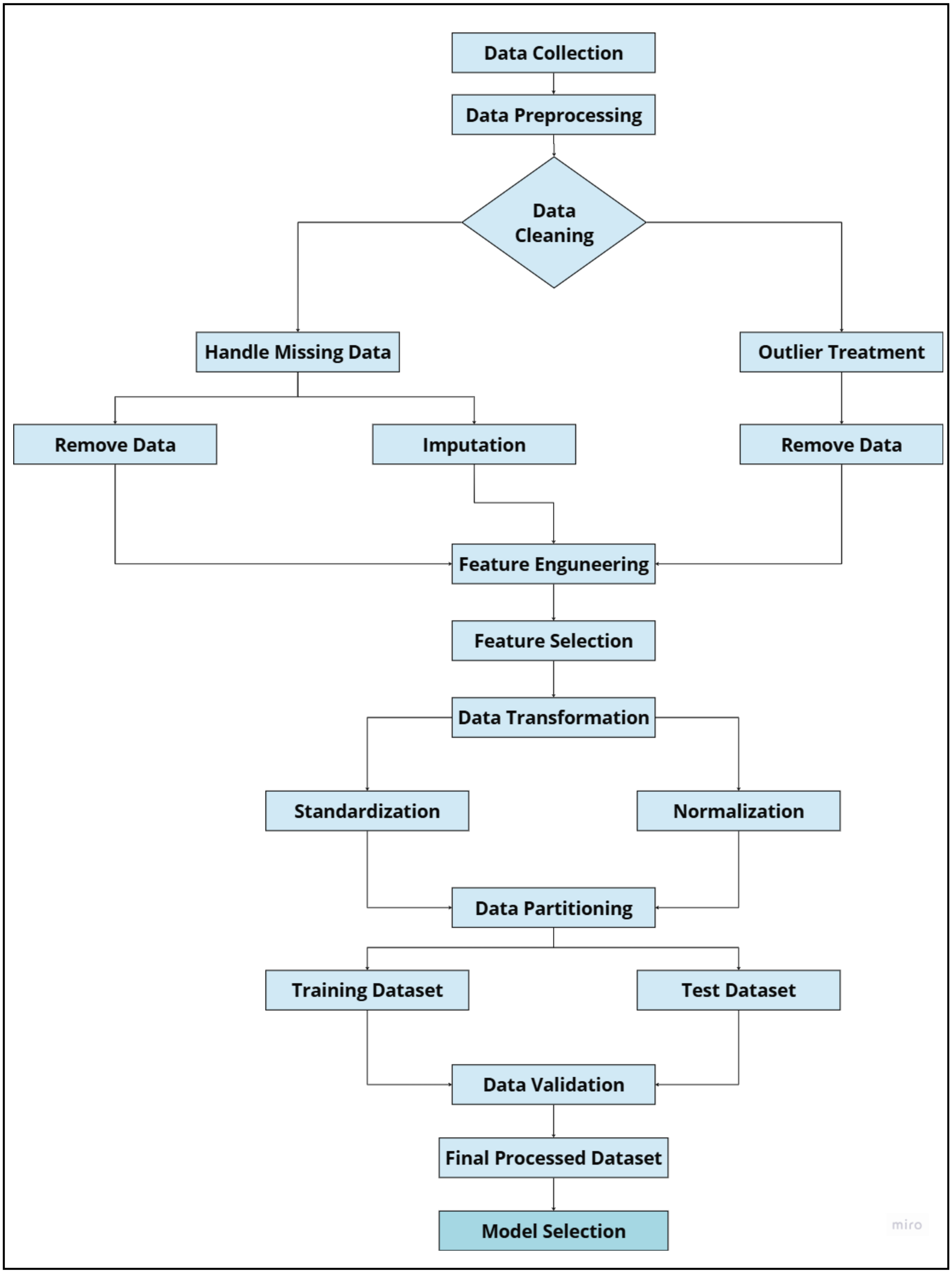

Figure 18: Flowchart Depicting the Preprocessing of Roll Grinding Production Data

4.4 Experimental Design

The experimental design of this thesis is structured to systematically evaluate the performance of various machine learning models in predicting optimal grinding parameters. This design ensures a robust and reliable comparison while addressing the research goals of minimizing errors and optimizing cycle times in grinding operations.

4.4.1 Overview of Experiment Goals

This research aims to improve grinding operations by using predictive models that minimize errors and maximize cycle times. It involves investigating the complex interactions that exist between grinding parameters such as depth of cut, wheel speed, and feed rate, and how these affect operational results like error rates and production efficiency. The presented work compares several of these in depth to select the most viable machine learning model for a task: classical, ensemble, and deep learning.

Investigate how fluctuations in the critical grinding parameters act upon the overall process and cycle times, as well as how they affect product quality and grinding efficiency. To identify the best method for predictive analytics, develop several machine learning models, train and evaluate them by comparing the results methodically. See how predictive analytics is really applied to grinding operations to predict the optimal parameter settings that will reduce production errors and enhance process efficiency.

4.4.2 Data Partitioning

It makes use of a number of very important partitioning and preprocessing of the data in training and testing datasets, critical steps which must be undertaken with due attention for any kind of machine learning process. As shown (Figure 22) a partition of the big dataset: from the historic production and unseen operationally into three mutually exclusive sets of training, validation, and test sets. This latter comprised 70% of the data, which trained the models to learn patterns and relationships intrinsic to it. This was also the important set during initial model development wherein various algorithms were first applied to iteratively arrive at better performance. The other subset representing 15% of the dataset was essential in fine-tuning the model's hyperparameters for an optimum fit to data. This subset allowed for the assessment of model performance during training and guided decisions on the adjustment of parameters with a view to optimizing the predictive capabilities of the models. Finally, the test set also constituted 15% of the data, which was withheld from the model during both training and validation. This holdout set was used only for the very final evaluation of the performance of the models in order to estimate unbiasedly how well the trained models generalize to unseen operational data.

Besides that, StandardScaler was applied to normalize the numerical features [56]. This scaling strategy changes data by normalizing each feature to a mean of zero and a standard deviation of one. Standardization is especially important in those models sensitive to the size of input data; these include models relying on the gradient-based optimization approaches-for example, neural networks, or distance metrics such as support vector machines or k-nearest neighbors. This strategy helped us to promote a more balanced model performance by ensuring that no particular feature, because of its bigger magnitude, dominated the learning process.

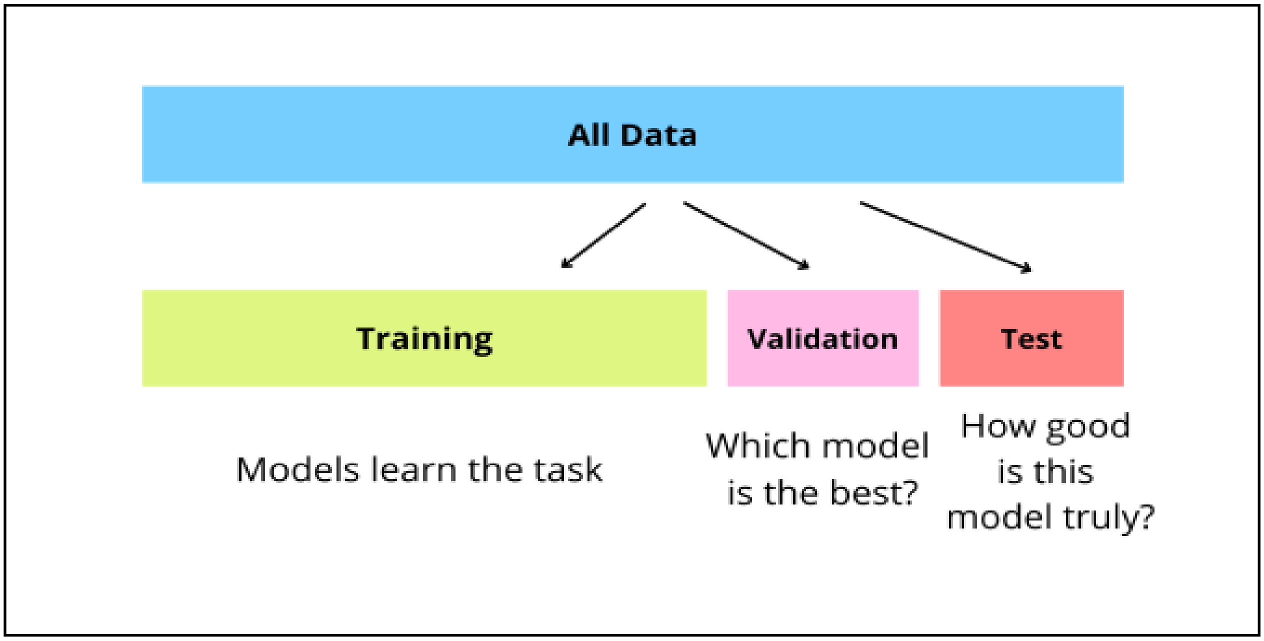

Figure 19: Data Partitioning: Training vs Validation vs Test [62]

Encoding for categorical features was done to transform the non-numeric values into a format useable by the machine learning algorithms. Label encoding and one-hot encoding were performed accordingly. Label encoding will be used for ordinal categorical variables while one-hot encoding will be used when there is no inherent order within the variable. These encoding strategies were applied to convert categorical variables into a form digestible to the machine learning algorithms, thereby making the learning process easier.

Together, these partitioning and preprocessing steps provided a solid foundation for model training, ensuring that the data was appropriately structured and normalized for optimal performance across the machine learning models.

4.5 Model Development

The model development process is range of machine learning models and also key steps such as feature selection, hyperparameter tuning, and model selection. These steps are all crucial with respect to improving performance and enhancing the generalization ability of the model.

4.5.1 Evaluation of Feature Selection

Feature selection is the process of finding and choosing the most pertinent features for model training; it is a very important preprocessing step in machine learning. The main criterion for feature selection in this thesis was mutual information. Mutual information is a measure of the dependency between two variables and provides the amount of knowledge that one variable offers about another. Some of its major uses are to find linear and nonlinear relationships between features and a target variable.

Mutual information was calculated for every feature concerning the target variable to measure the strength of dependencies. Then, features with high mutual information with the target variable were kept and passed for training the model, whereas weak or negligible dependencies were thrown out. It was also supposed to reduce the dimensionality of the dataset by removing the redundant or irrelevant features that could introduce noise into the model and degrade the performance.

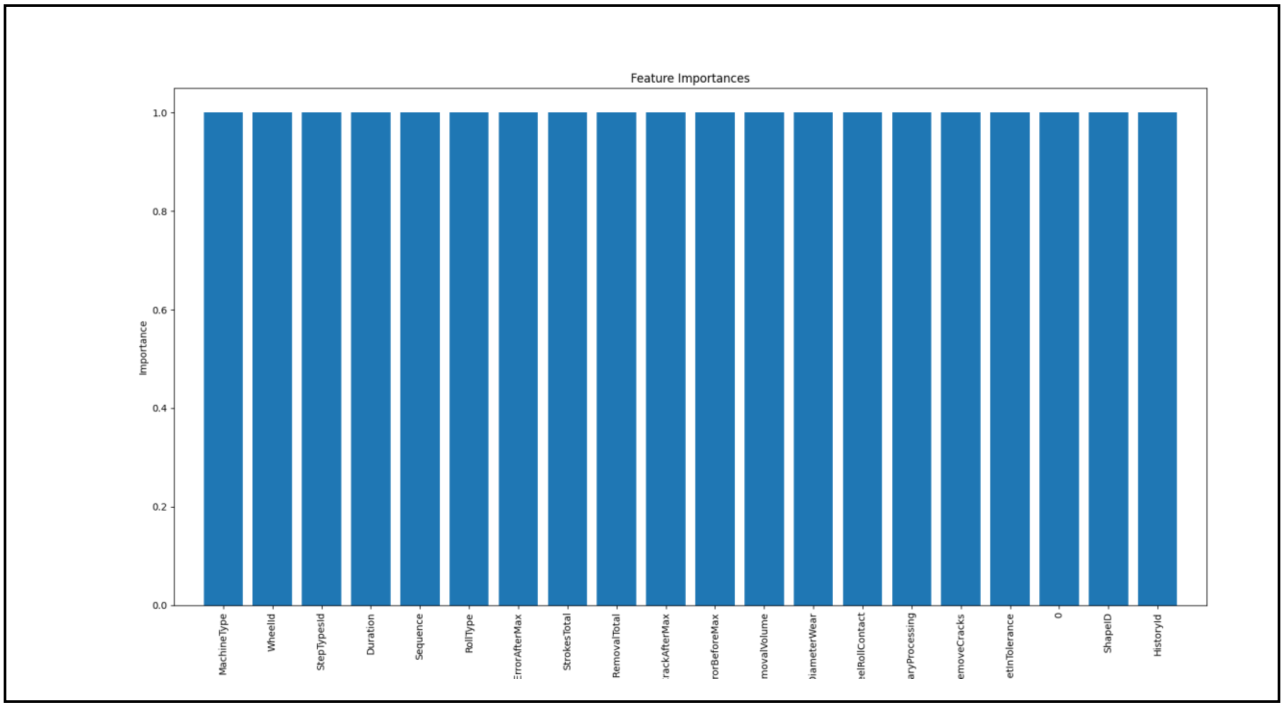

Figure 20: Feature Selection using Mutual Information and Select only Relevant Feature

As shown (Figure 20), selected based on mutual information, had high dependency scores. Hence holding the most relevant information for any predictive task. This has further resulted in lighter models with improved accuracy and computation time because these models have been trained on a reduced yet more informative feature set.

Further, the mutual information-based feature selection process was visualized in graph form to show the strength of the dependency existing between each feature and the target variable. Features with the highest dependency scores are clearly highlighted in order to inform model development. This is represented visually in the image below, showing the ranking of features by their mutual information values. These types of diagrams are immensely useful to designers for understanding the reasoning behind feature selection and ensuring that the most informative features are retained.

4.5.2 Hyperparameter Tuning

Hyperparameter tuning became the next most important phase in model development after the feature selection process [57]. In simple terms, hyperparameters are crucial settings that dictate how a machine learning model will be trained. Examples include the learning rate, the number of trees for a random forest, the depth of decision trees, and the number of layers or units for deep learning models. Selection of hyperparameters is an important task since hyperparameters are the most influential factors on the performance of the model. Poor hyperparameter choice may result in overfitting, where the model memorizes the training data and performs badly on unseen data, or underfitting, where models are too simple to capture the underlying patterns in the data.

In this work, we have used Grid Search to tune the hyperparameters, a rather exhaustive but efficient approach in determining the best hyperparameters that should serve a particular model. Grid search works by exhaustively searching through a manually specified set of hyperparameter values. For each combination of hyperparameters, the model is trained and evaluated, and the performance metrics are recorded. This approach will ensure that every combination within the pre-defined hyperparameter space is tried out, thus enabling a structured approach towards model optimization.

For each model, a set of hyperparameters was selected based on domain knowledge and typical values used for similar tasks. Subsequently, different configurations of these hyperparameters were explored using the grid search algorithm. For example, in the Random Forest model, the number of trees (estimators) and maximum tree depth were varied, while for SVM, the regularization parameter and kernel type were tested. The grid search was conducted using the validation set, which allowed the model's performance to be evaluated under each configuration. This was meant to fine-tune the hyperparameters and try to get the best out of the model from data that it hadn't seen during training.

Later, after running grid search and pinpointing the optimal hyperparameters, the model was retrained on the entire training dataset using the best hyperparameters identified. This will make sure that the final model is fully optimized and can make as accurate predictions as possible when evaluated on unseen data. By using grid search for hyperparameter tuning, we made sure that each model was carefully optimized, hence improving the model's accuracy and robustness while minimizing the risk of overfitting or underfitting.

The use of grid search for hyperparameter tuning is beneficial mainly for its completeness, which can allow for the most optimal configuration of hyperparameters. It is computationally expensive and usually used where large datasets and complex models are involved. Despite that, improvements in accuracy and model performance from grid search present it as an important technique to have the best possible configuration for each different machine learning model.

4.5.3 Model Selection

Model selection was the last step of the model development process, which is about the evaluation of various machine learning models and choosing the one that gave the best generalization performance. It was required to compare multiple models, ranging from traditional models to ensemble methods and deep learning models, after they were trained and fine-tuned.

The selection criteria ranged from quantitative measures, including accuracy, precision, recall, F1-score, and AUC, to qualitative ones, such as model interpretability and training time [58]. The performance of the models was evaluated on the test set, which had been kept separate from the training and validation sets to ensure that the evaluation was unbiased and reflective of the model's ability to generalize to unseen data.

The focus was on deploying superior-performing models. If performance was inadequate, consideration was made for custom ensembles where the strengths of various models could be combined to produce a better overall performance than one alone would. Select the best model that can provide the most robust, accurate, and interpretable results for final use.

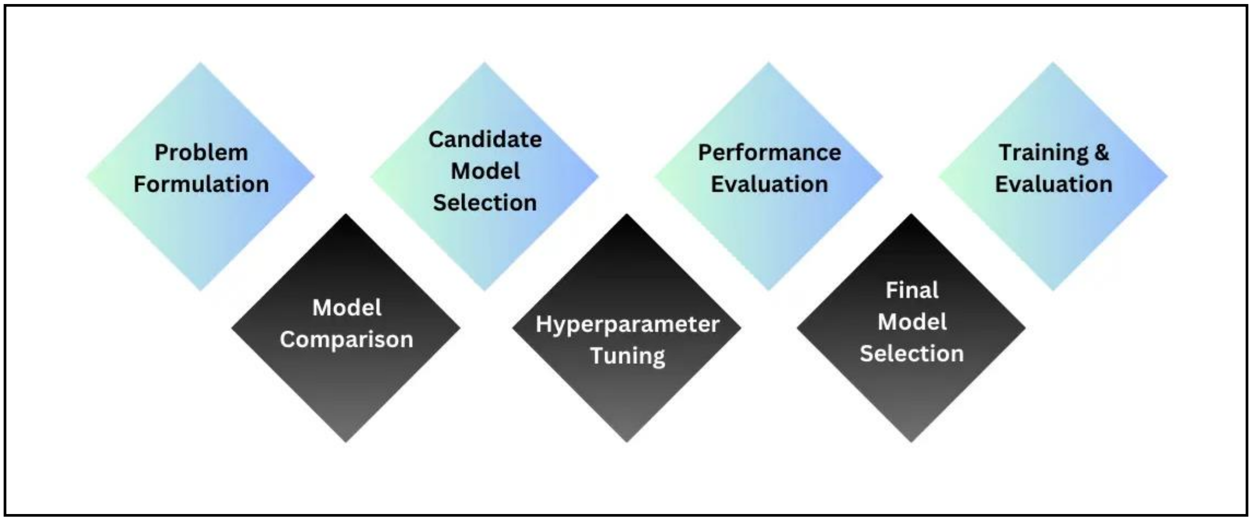

Figure 21: Process of Model Selection [58]

It is this organized process of model development-from the choice of feature pre-processing through the steps of hyperparameter tuning and thoughtful model selection-that had made the exploration effective in landing at the best performance by a machine learning model for the dataset being considered. All of these steps mean that, in addition to the high predictive accuracy being met, the chosen model will be low-cost, interpretable, and generalizable.

5. Results

This chapter describes a set of experiments in order to study the prediction and analysis of main parameters of roll grinding machines using different machine learning algorithms. The main goal of these experiments was to establish up to what extent such models are capable of predicting critical operational parameters, which influence directly the quality and effectiveness of the grinding process. Advanced algorithms will be utilized to seek out patterns and relationships that exist between the machine settings-like grinding speed, feed rate, grinding force, coolant flow, among others-and the resultant surface quality of the rolls being ground.

The modeling was done with a combination of traditional classifiers such as Random Forest and K-Nearest Neighbors, advanced ensemble models including XGBoost, GBR, and LightGBM, and other methods like Long Short-Term Memory networks. The various models were compared in terms of their predictive capability and their ability to generalize across the operating conditions for the grinding process. These models were evaluated based on F1-score and accuracy with a confusion matrix, which presents a comprehensive result about the strengths and weaknesses of the models under consideration.

5.1 Result of Machine Learning Algorithms

The results of applying various machine learning algorithms for this thesis yielded insightful findings. By evaluating models like Random Forest, K-Nearest Neighbors, advanced ensemble models including XGBoost, GBR, and LightGBM, and other methods like Long Short-Term Memory networks, a comprehensive analysis of their performance was conducted. The outcomes showcased the strengths and weaknesses of each algorithm.

5.1.1 Result of Random Forest (RF)

The key parameters of the roll grinding machine were predicted using the Random Forest algorithm. This ensemble method, which builds multiple decision trees and combines their predictions, can be used for complex datasets and also helps in finding non-linear relationships between input features and target variables. Trying to improve the performance of this model, hyperparameter tuning has been done using RandomizedSearchCV—a method that systematically tries a predefined range of parameters to find the best combination.

Table 1: The Hyperparameter Search Space for RF Model

Number of estimators (n_estimators)	[100, 200, 300, 400]
Maximum features (max_features)	['auto', 'sqrt', 'log2']
Maximum depth (max_depth)	[10, 20, 30, None]
Minimum samples split (min_samples_split)	[2, 5, 10]
Minimum samples leaf (min_samples_leaf)	[1, 2, 4]
Bootstrap sampling (bootstrap)	[True, False]

Table 2: The Best Hyperparameters for the RF Model

n_estimators	200
min_samples_split	2
min_samples_leaf	4
max_features	Log2
max_depth	None
bootstrap	True

With this parameter, the Mean Squared Error (MSE) is 16.79, whereas the Coefficient of Determination (R^2) is 0.255.

Furthermore, the results of confusion matrix are presented in Figure 22.

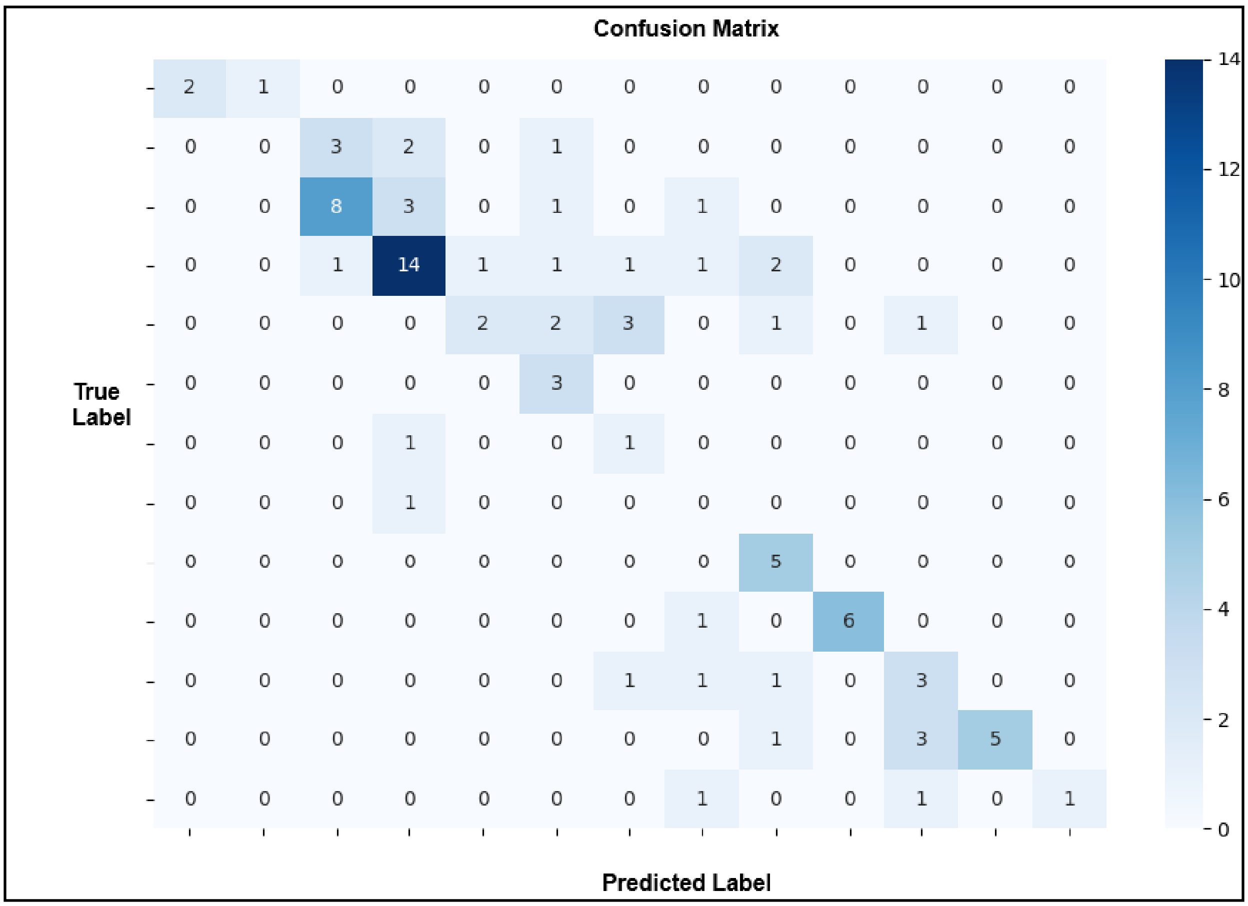

Figure 22: Confusion Matrix for RF Model

The RF model's moderate performance, as indicated by the evaluation metrics, highlights its limitations as a standalone approach for predicting roll grinding machine parameters. Even though the hyperparameter tuning improved the model a bit, the results indicate that the RF model has difficulty understanding all the complex relationships in the dataset. While it generalizes well across different operating conditions, its predictive power is still not good enough for applications requiring high accuracy and reliability. Given these limitations, there is a need to investigate alternative models that may offer improved performance.

5.1.2 Result of K-Nearest Neighbors (KNN)

The K-Nearest Neighbors model is applied for the prediction of the parameters of a roll grinding machine. KNN is one of the simple, interpretable algorithms that make predictions based on similarities between data points. The KNN model was tuned by a grid search to optimize the hyperparameters of the algorithm. The grid of parameters involves a variation in the number of neighbors, distance metrics, and weighting schemes for predictions.

Table 3: The Hyperparameters for KNN Model

Number of Neighbors (n_neighbors)	[3, 5, 7, 10]
Distance Metric (p)	[1, 2]
Weights	['uniform', 'distance']

Table 4: The Best Hyperparameters for the KNN Model

n_neighbors	10
Distance Metric (p)	1 (Manhattan distance)
Weights	Uniform

The KNN model's performance was assessed using key evaluation metrics, which are summarized, Mean Squared Error (MSE) is 49.11 and R^2 Score is -0.11. The results of Actual vs Predicted Values are presented in Figure 23.

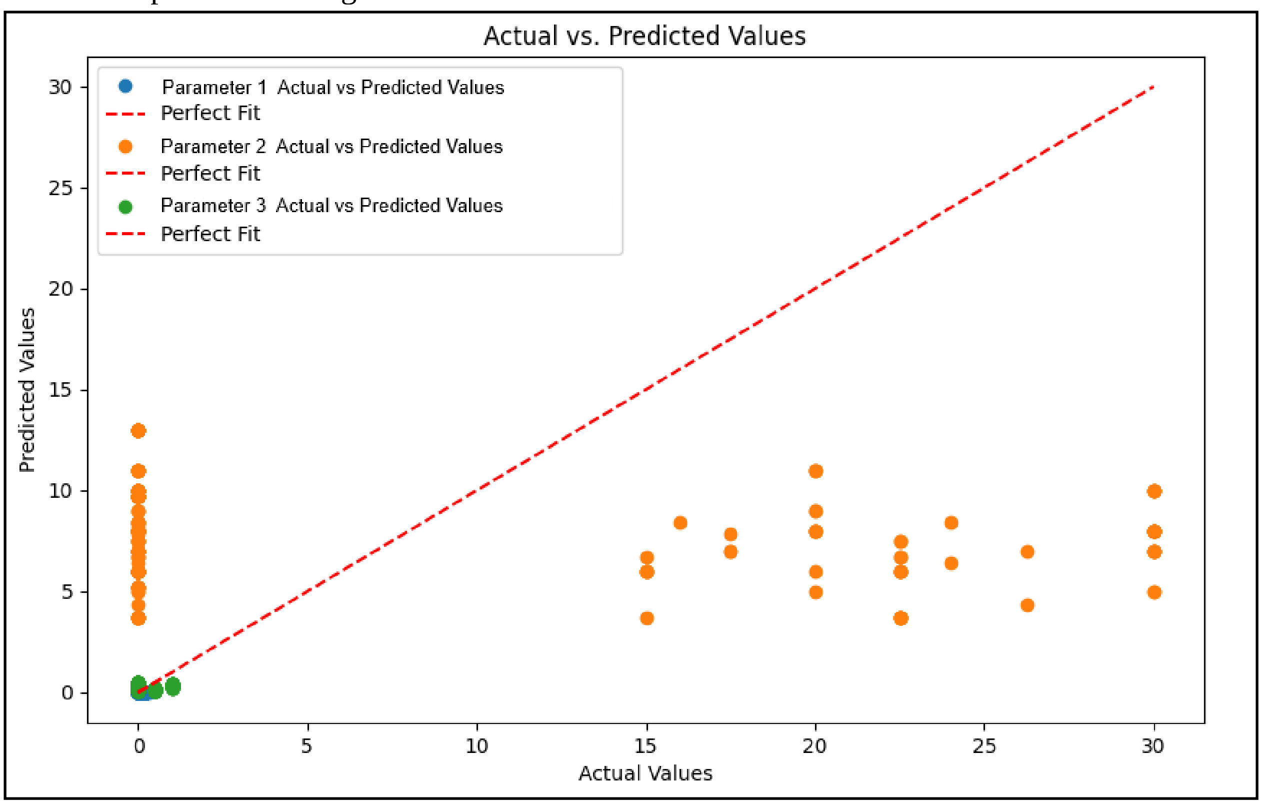

Figure 23: Actual vs Predicted Values of KNN Model

The results indicated that the KNN model was unsatisfactory in terms of predictive accuracy on the dataset, as shown by the high MSE, coupled with a negative R^2 score. This result may suggest that the model did not capture the relationships underlying the different input features and target variables well. Several factors likely contributed to this outcome, including the high dimensionality of the dataset, which diminishes the effectiveness of the distance metric, and the non-linear nature of the relationships among roll grinding parameters, which KNN's reliance on local averaging cannot adequately address. Moreover, the sensitivity to outliers and noisy data could have further impaired the predictive performance of KNN.

This underperformance highlights the need to explore alternative machine learning approaches better suited to handling complex, non-linear relationships. Ensemble models, such as Gradient Boosting Regressor (GBR) and Extreme Gradient Boosting (XGBoost), are particularly promising due to their ability to model intricate patterns and interactions in the data.

5.1.3 Result of Gradient Boosting Regressor (GBR)

The Gradient Boosting Regressor (GBR) model was applied to predict the parameter of operations in the roll grinding machine process. The model was optimized using RandomizedSearchCV, which enabled the identification of the best-performing hyperparameters.

Table 5: The Hyperparameters for the GBR Model

Number of boosting stages (n_estimators)	[100, 200, 300, 400]
Learning rate	[0.01, 0.05, 0.1, 0.2]
The maximum depth of the trees (max_depth)	[3, 4, 5, 6]
The minimum number of samples to split (min_samples_split)	[2, 5, 10]
The minimum number of samples to be at leaf (min_samples_leaf)	[1, 3, 5]
Subsample	[0.8, 0.9, 1.0]

The following hyperparameters were found to provide the best results for the prediction task.

Table 6: The Best Hyperparameters for the GBR Model

subsample	0.9
n_estimators	400
min_samples_split	5
min_samples_leaf	5
max_depth	4
learning_rate	0.01

These optimized hyperparameters allow the model to generalize well, balancing model complexity with avoiding overfitting. The performance of the GBR model was evaluated using key regression metrics such as the Mean Squared Error (MSE) and the R^2 score. The results were as follows, Optimized Mean Squared Error (MSE) is 16.7679 and R^2 Score is 0.2562. the MSE indicates that the GBR model performs moderately well, but there is room for improvement in prediction accuracy. The R^2 score of 0.2562 indicates that approximately 25.62% of the variance in the target variable is explained by the model. This suggests that the GBR model captures some of the underlying patterns in the dataset but that other factors may also influence the sequence that are not fully accounted for by the model.

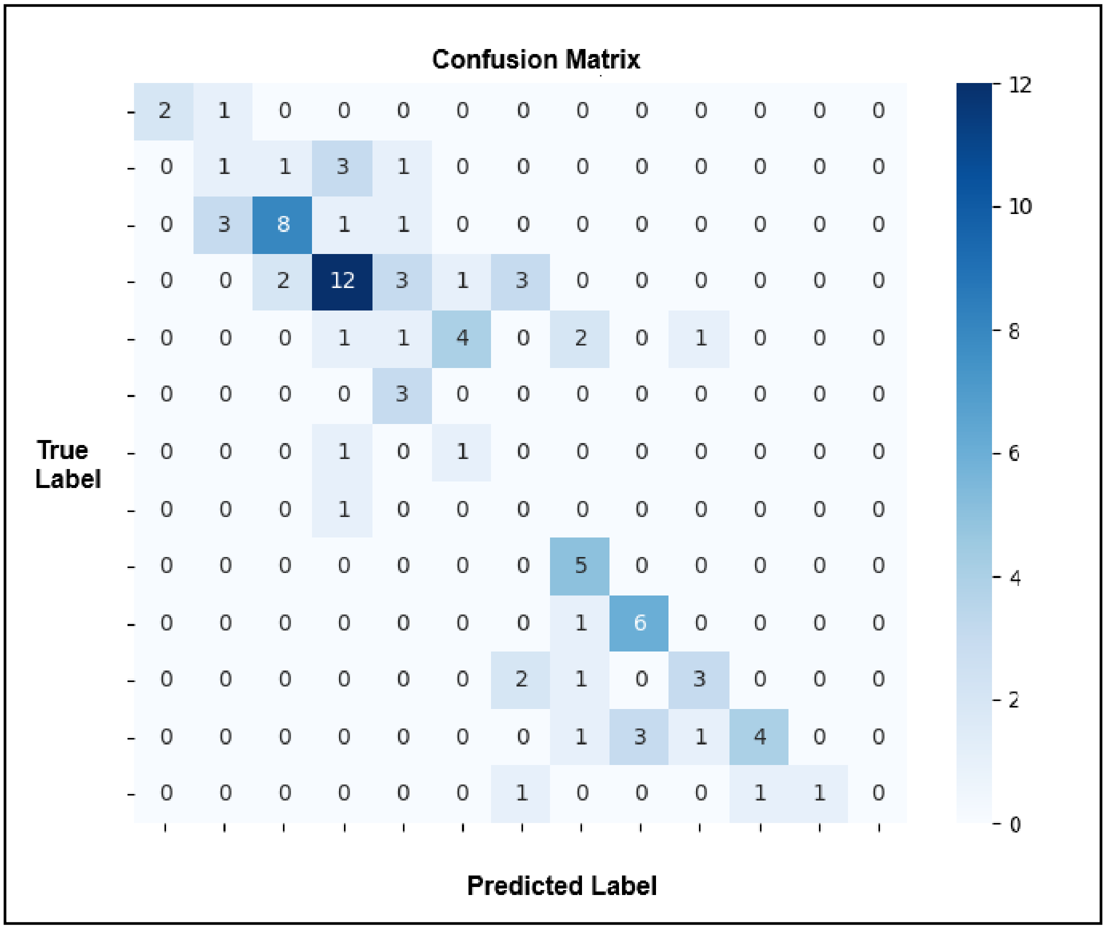

Figure 24: Confusion Matrix for GBR Model

Since GBR models are powerful in modeling nonlinear relationships, their use in predicting operational parameter of roll grinding machines is promising. However, achieving higher predictive accuracy requires an advanced model.

5.1.4 Result of Extreme Gradient Boosting (XGBoost)

The XGBoost model was implemented as part of the predictive framework to model the roll grinding machine parameters. XGBoost is one of the most efficient and working models, mainly for large-size data with complicated relations between the variables. The model was fine-tuned using RandomizedSearchCV, which allowed for an optimal combination of hyperparameters.

Table 7: The Hyperparameters for the XGBoost Model

Number of estimators (n_estimators)	[100, 200, 300, 400]
Learning rate (learning_rate)	[0.01, 0.05, 0.1, 0.2]
Maximum depth (max_depth)	[3, 4, 5, 6]
Minimum child weight (min_child_weight)	[1, 3, 5]
Subsample ratio (subsample)	[0.8, 0.9, 1.0]
Column sample ratio per tree (colsample_bytree)	[0.8, 0.9, 1.0]

In following table 8 were found the best hyperparameter for the XGBoost model.

Table 8: The Best Hyperparameters for the XGBoost Model

n_estimators	100
learning_rate	0.05
max_depth	5
min_child_weight	5
subsample	0.9
colsample_bytree	0.9

Using the optimized parameters, the XGBoost model achieved an accuracy of 43.18%, indicating that it correctly classified 43.18% of the samples. The model had an F1 score of 0.42, hence a moderately balance between precision and recall. Specifically, the recall was 43.18%, signifying that the model identified 43.18% of the actual positive cases, while the precision was 44.34%, highlighting that 44.34% of the predicted positive cases were correct. These results reflect the model's moderate performance, hence suggesting room for optimization in order to improve the classification results.

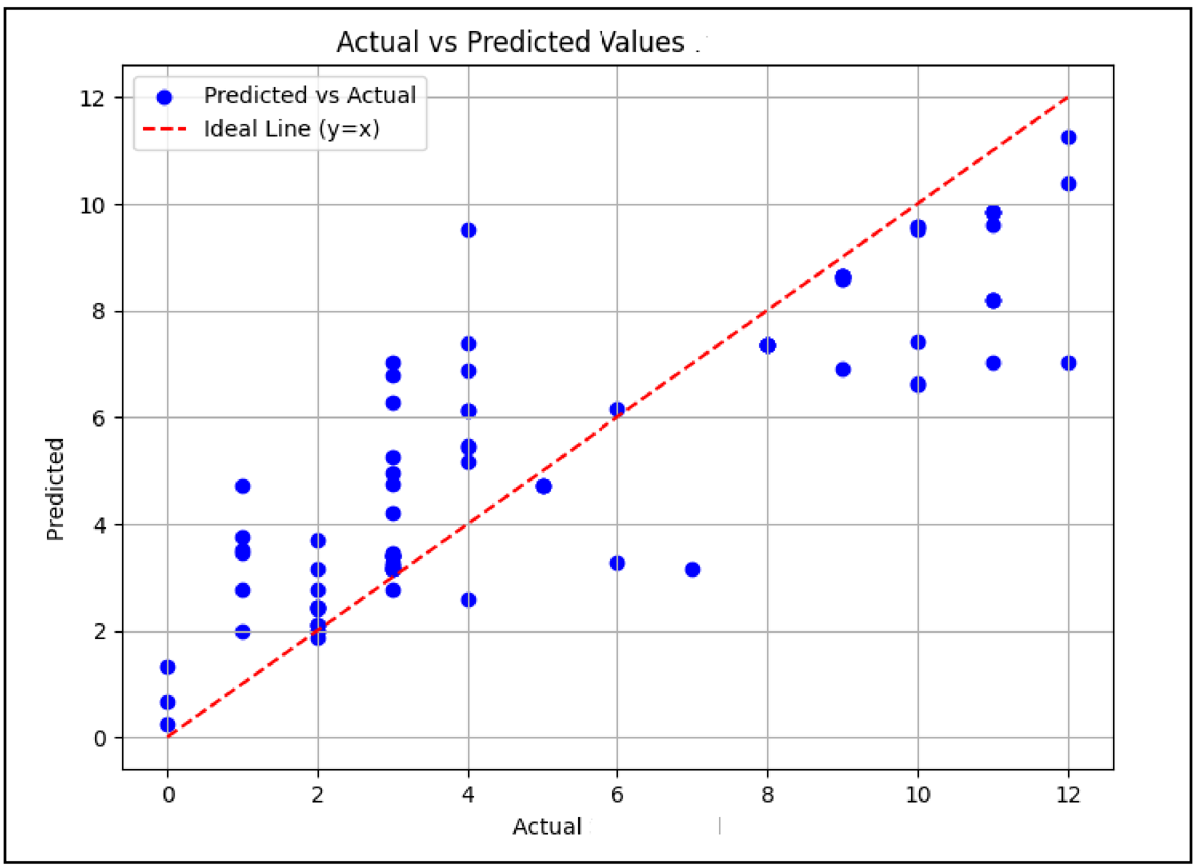

Figure 25: Actual vs Predicted Values of XGBoost Model

Figure 27 presents a graphical comparison of actual versus predicted values for further insights into the performance of the model. Although the plot shows some positions where the model correctly predicted the actual class, there are also some discrepancies between actual and predicted values.

These discrepancies illustrate the model's struggles to generalize patterns in the data, contributing to its moderate overall accuracy. These results imply that factors such as dataset complexity and possible class imbalances may have impacted performance.

5.1.5 Result of Ensemble Model (Support Vector Machines (SVM) + Random Forest)

The ensemble model combining Support Vector Machine (SVM) and Random Forest (RF) was implemented to leverage the strengths of both algorithms and improve classification accuracy. SVM is particularly effective for high-dimensional data and excellent at finding the optimal hyperplane for classification, while Random Forest is robust to noise and captures complex interactions among features. A soft voting strategy was employed in this ensemble, allowing the model to weigh predictions probabilistically.

In table 9 were found the hyperparameter for the SVM model.

Table 9: The Hyperparameters for the SVM Model

Regularization parameter(C)	[0.1, 1, 10, 100]
Kernel coefficient(gamma)	['scale', 'auto']
Kernel type (kernel)	['linear', 'rbf', 'poly', 'sigmoid']

Table 10: The Best Hyperparameters for the SVM Model

C	0.1
gamma	Scale
kernel	Linear

Default parameters were used for Random Forest during the ensemble creation, as the primary focus was on integrating the two models in a voting classifier. The performance of the ensemble model was evaluated using various metrics, including the confusion matrix, classification report, and cross-validation results.

The confusion matrix showed a major problem in misclassifications within the 12 classes, which pinpointed the inability of the model to distinguish between at least two classes. The classification report had given an average precision score of 0.59, a recall score of 0.45, and an F1-score of 0.66 across the classes. There is evident a huge trade-off between precision and recall, meaning that it can identify some classes quite well while failing completely in certain others. The overall accuracy of the model stood at 68%, indicating that the ensemble model correctly classified 68% of the samples. Besides, the cross-validation results showed that the model performance varied, with a mean accuracy of 68% and a standard deviation greater than zero, reflecting inconsistent outcomes across different folds. Collectively, these metrics indicate the ensemble model's moderate performance and point to areas for improvement to further develop the model's capability to handle the multi-class classification problem.

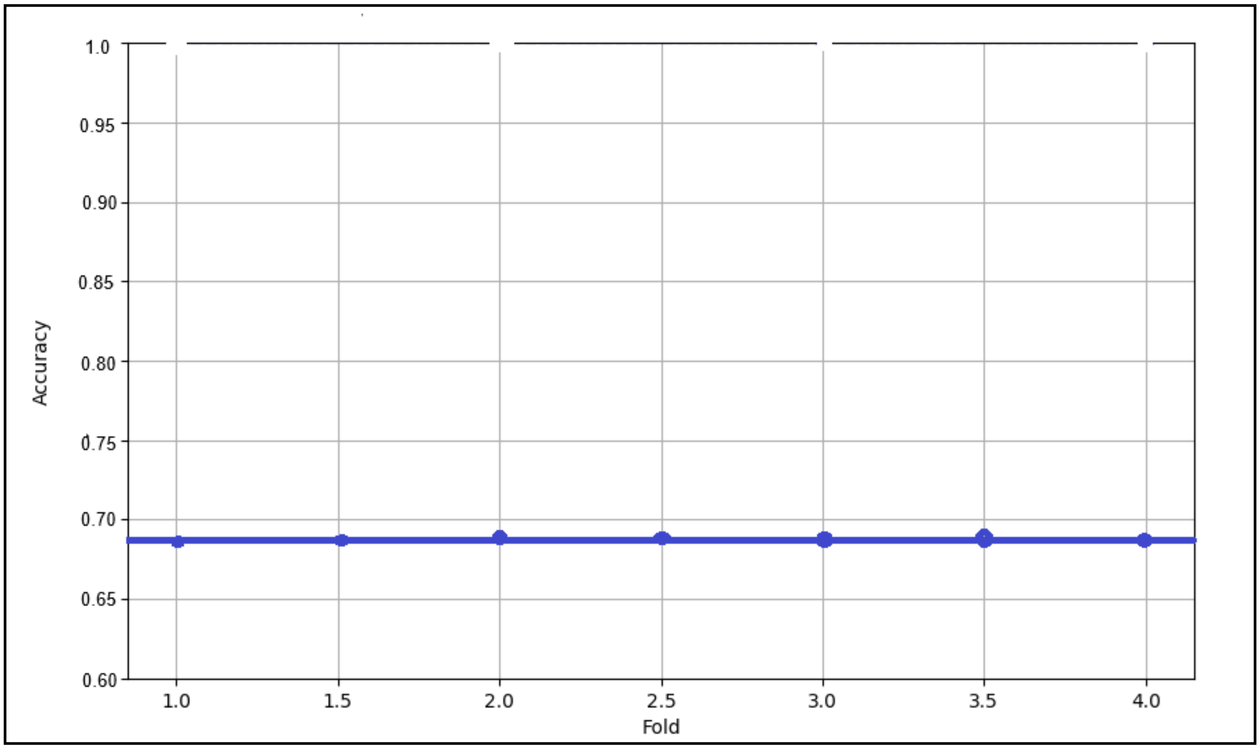

Figure 26: Overall Accuracy of (SVM + RF) Ensemble Model

5.1.6 Result of Ensemble Model (GBR + XGBoost)

The ensemble model utilized both the Gradient Boosting Regressor (GBR**)** and the XGBRegressor. Each component underwent optimization using RandomizedSearchCV to identify the best hyperparameters for enhanced performance.

Table 11: The Hyperparameters for the Ensemble (GBR) Model

Number of boosting stages (n_estimators)	[100, 200, 300, 400]
learning_rate	[0.01, 0.05, 0.1, 0.2]
Maximum depth of the individual regression (max_depth)	[3, 4, 5, 6]
Minimum number of samples to split (min_samples_split)	[2, 5, 10]
Minimum number of samples to be at a leaf (min_samples_leaf)	[1, 3, 5]
Fraction of samples used for fitting subsample	[0.8, 0.9, 1.0]

Table 12: The Hyperparameters for the Ensemble (XGBoost) Model

Number of boosting stages (n_estimators)	[100, 200, 300, 400]
learning_rate	[0.01, 0.05, 0.1, 0.2]
Maximum depth of the individual regression (max_depth)	[3, 4, 5, 6]
Minimum sum of instance weights (min_child_weight)	[1, 3, 5]
Proportion of the training data (subsample)	[0.8, 0.9, 1.0]
Proportion of features (columns) (colsample_bytree)	[0.8, 0.9, 1.0]
Minimum loss reduction (gamma)	[0, 0.1, 0.2, 0.3]

The best hyperparameter for ensemble model is following,

Table 13: The Best Hyperparameters for the Ensemble (GBR) Model

subsample	0.8
n_estimators	400
min_samples_split	10
min_samples_leaf	5
max_depth	6
learning_rate	0.01

Table 14: The Best Hyperparameters for the Ensemble (XGBoost) Model

subsample	0.8
n_estimators	300
min_child_weight	3
max_depth	6
learning_rate	0.01
gamma	0.1
colsample_bytree	1.0

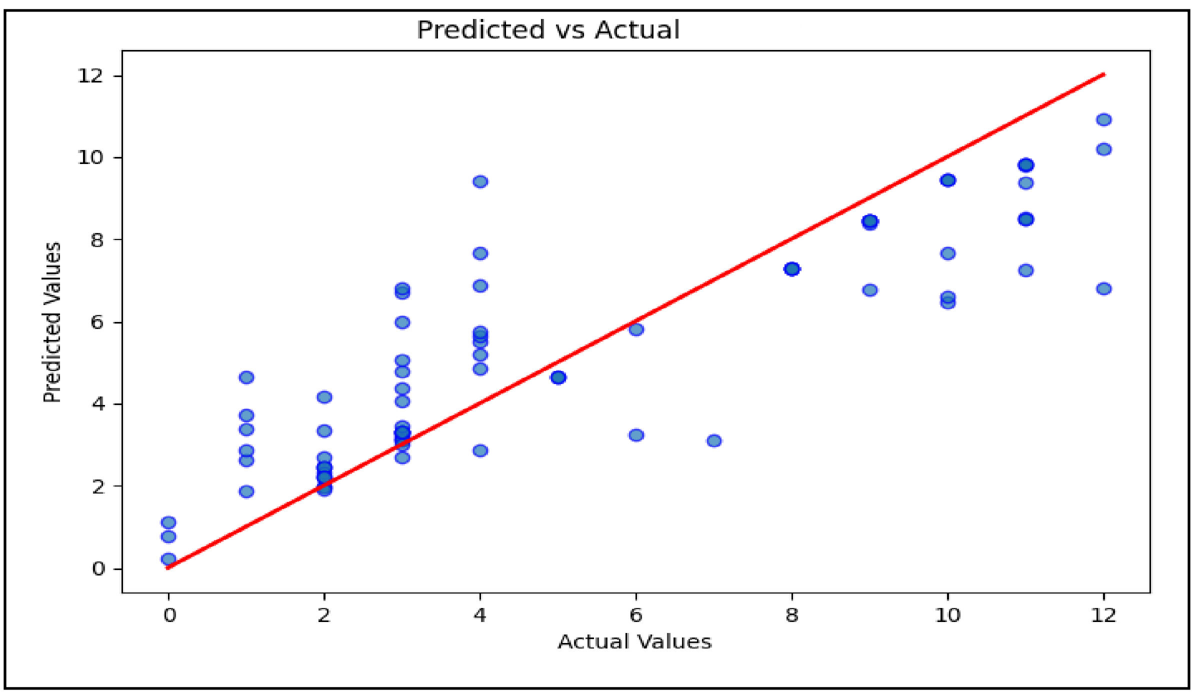

Figure 27: Actual vs Predicted Values of Ensemble (GBR + XGBoost) Model

The predictive performance of the ensemble model was checked on a set of regression metrics that indicated how well it would work in predicting the target variable. Whereas the individual cross-validated MSE for GBR and XGBRegressor models was 6.79, for the ensemble model, this reduced to a value of 3.29. This substantial improvement underlines the synergistic advantage of integrating both regressors within an ensemble framework. The R-squared value of the model was 0.75, indicating that it explains 75% of the variance in the target variable, reflecting its strong ability to capture underlying relationships within the dataset. The MSE was calculated at 1.31, which signifies that, on average, the model's predictions deviated from the actual values by 1.31 units. Besides, the root mean squared error was 1.81, further emphasizing the efficiency of the model in minimizing the larger errors since RMSE punishes the larger deviations much harder than MAE. These results thus show the substantial predictive power and error rate reduction ability of the ensemble model compared to that of the single models. Low values of MSE and RMSE and a high R-squared value support this being a very strong ensemble method, which performs much better when there is more intricate and nonlinear behavior in the data.

5.1.7 Result of Ensemble Model (XGBoost + Categorical Boosting (CatBoost) + Support Vector Regression (SVR) + LightGBM)

The ensemble model that combines various advanced machine learning algorithms, including XGBoost, LightGBM, SVR, and CatBoost, was evaluated to predict the target variable. Below are the best hyperparameters found for each individual model after performing a randomized search:

Table 15: The Best Hyperparameters for the Ensemble (XGBoost) Model

n_estimators	learning_rate	max_depth	subsample	colsample_bytree
200	0.01	5	0.9	0.8

Table 16: The Best Hyperparameters for the Ensemble (LightGBM) Model

n_estimators	learning_rate	num_leaves	min_child_samples	subsample
400	0.05	70	10	0.8

Table 17: The Best Hyperparameters for the Ensemble (SVR) Model

C	epsilon	gamma
10	0.2	scale

Table 18: The Best Hyperparameters for the Ensemble (CatBoost) Model

iterations	learning_rate	depth	subsample
300	0.01	6	0.9

The model's performance was assessed using several common regression metrics, including Mean Squared Error (MSE), R-squared (R^2), and Root Mean Squared Error (RMSE). The calculated Mean Squared Error stands at 5.32, meaning the difference between actual values and predicted values on average, from this model, is 5.32 squared units. Although not a really small value, it stays in a decent range considering the nature of the task-pretty complex regression. The R-squared value of 0.72 suggests that the model explains 72% of the variance in the target variable. This suggests the model is getting a fair amount of information from the variability in the dataset, although considerable room for further refinement of the model can be expected. The Root Mean Squared Error (RMSE) was 1.99, highlighting the model's ability to reduce larger errors, with RMSE penalizing significant deviations more heavily. While the RMSE value is reasonable, it also shows that there are some fairly large prediction errors that could be improved by subsequent models.

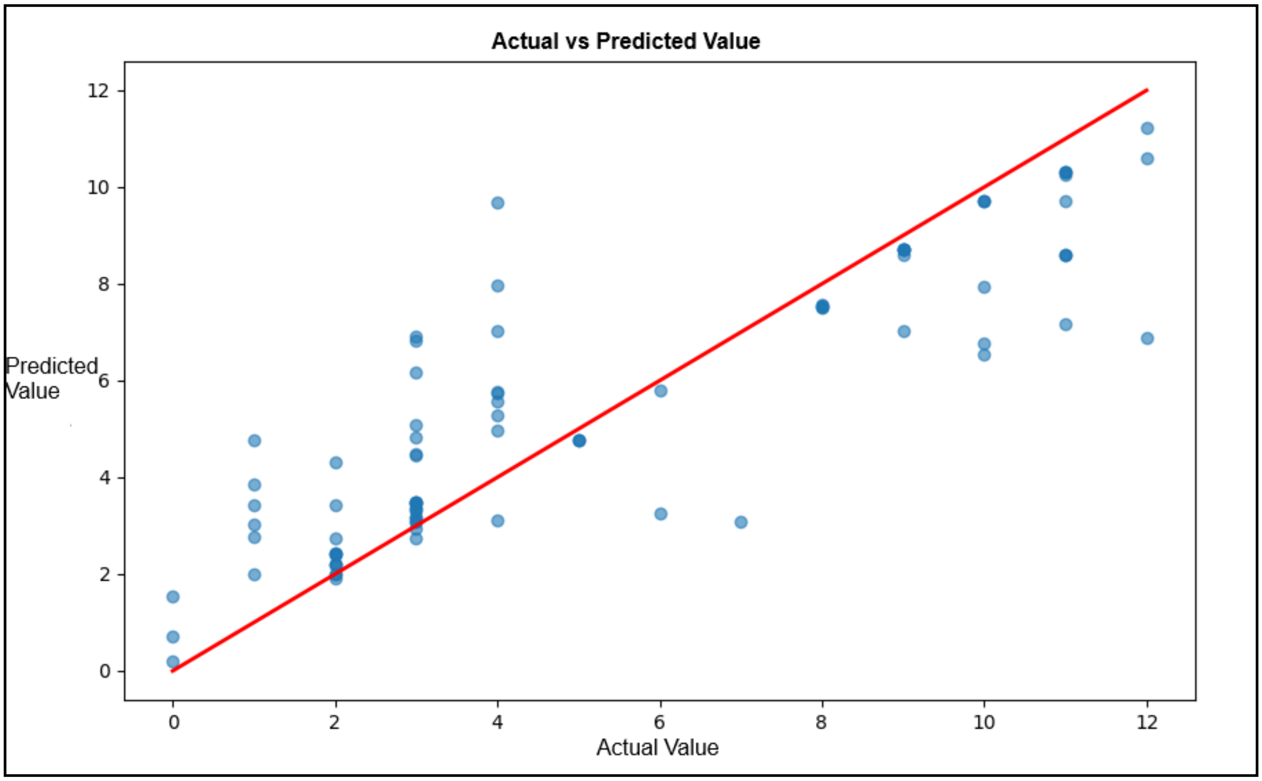

Figure 28: Actual vs Predicted Values of Ensemble (XGBoost + LightGBM + SVR + CatBoost) Model

Although the ensemble model, which combines XGBoost, LightGBM, SVR, and CatBoost, demonstrated moderate predictive accuracy, it did not achieve the high level of performance. The obtained MSE and RMSE quantities show there is still some level of error in the model's predictions, while the R^2 score of 0.72 suggests that the model explains a good portion of the variance but leaves some room for further refinement.

5.1.8 Result of Long Short-Term Memory (LSTM)

So far, apart from all the mode, the performance of LSTM is best with the hyperparameters given below:

Table 19: The Hyperparameters for the LSTM Model

Number of units (neurons) (units)	[32, 64]
The input units to drop (dropout_rate)	[0.1, 0.2, 0.3]
learning rate (learning_rate)	[0.001, 0.01]
Number of samples processed before updates weights (batch_size)	[16, 32]

Table 20: The Best Hyperparameters for the LSTM Model

Number of units (neurons) (units)	64
The input units to drop (dropout_rate)	0.1
learning rate (learning_rate)	0.01
Number of samples processed before updates weights (batch_size)	16

The Long Short-Term Memory (LSTM) model, optimized with hyperparameter tuning, showed as the most accurate predictive model in this thesis. Using the parameter configuration of 64 units, a dropout rate of 0.1, a learning rate of 0.01, and a batch size of 16, the model demonstrated superior performance across multiple evaluation metrics.

The MAE for this model was 0.00485, reflecting a very small average difference between the actual values and the predictions of the model. The Mean Squared Error (MSE) was recorded at 0.0000325, a remarkably low value reflecting the model's precision in minimizing error variance. The RMSE was similarly low at 0.00570, which hence underlines the capability of this model to deal with deviations on outliers effectively. The R-squared (R^2) value was an outstanding 0.9995, indicating that the LSTM model explains 99.95% of the variance in the target variable. The high R^2 underlines that this model is well fitted and reliable to catch the complex temporal relations within this dataset.

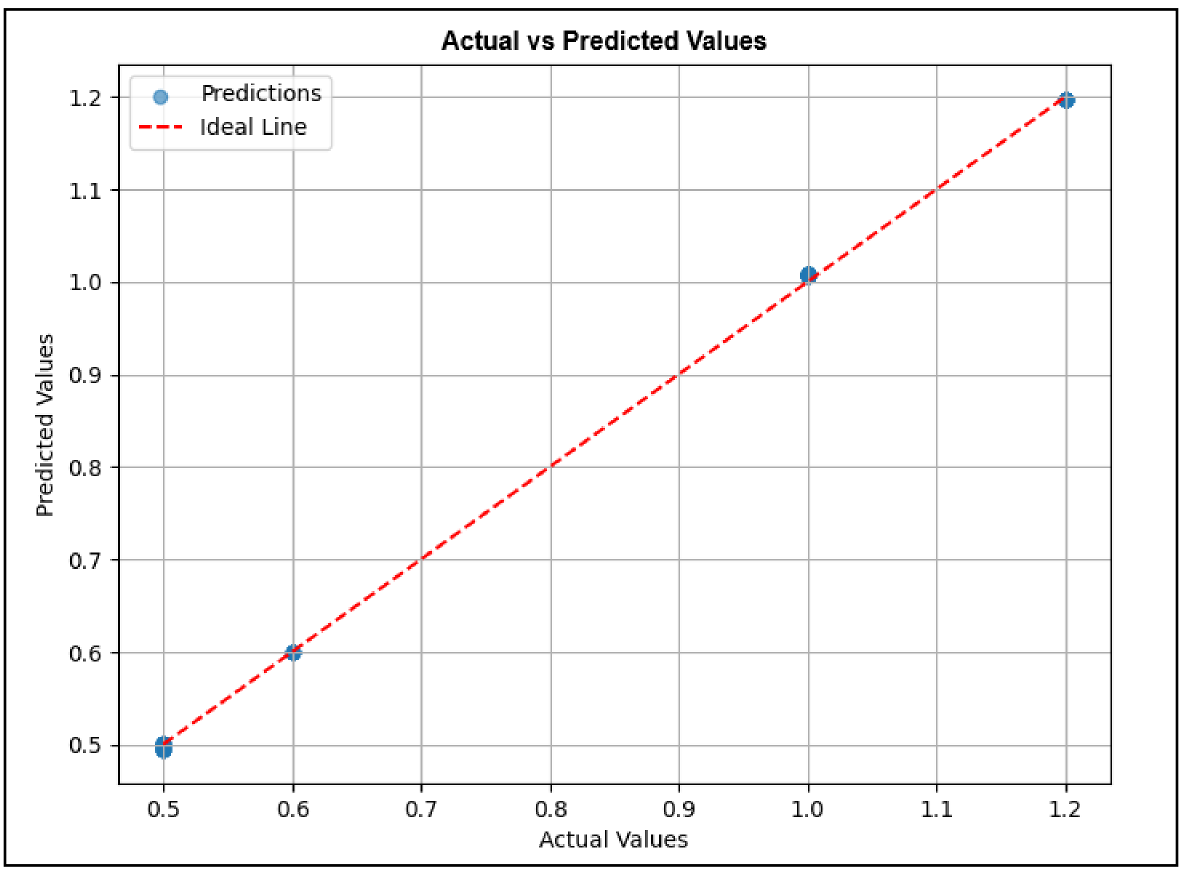

Figure 29: Actual vs Predicted Values of LSTM Model

Figure 31, comparing actual versus predicted values is presented to show the very good predictive accuracy of the LSTM model. The predicted values align closely with the actual values, forming a near-perfect linear relationship. This shows that the generalization achieved by the model is very good and deliver accurate predictions with minimal deviation.

The superior performance of the model is further supported by the smooth convergence of the training and validation loss curves, there is not excessive overfitting or underfitting during training. This indicates that the model effectively learned from the data without succumbing to noise or over-reliance on training.

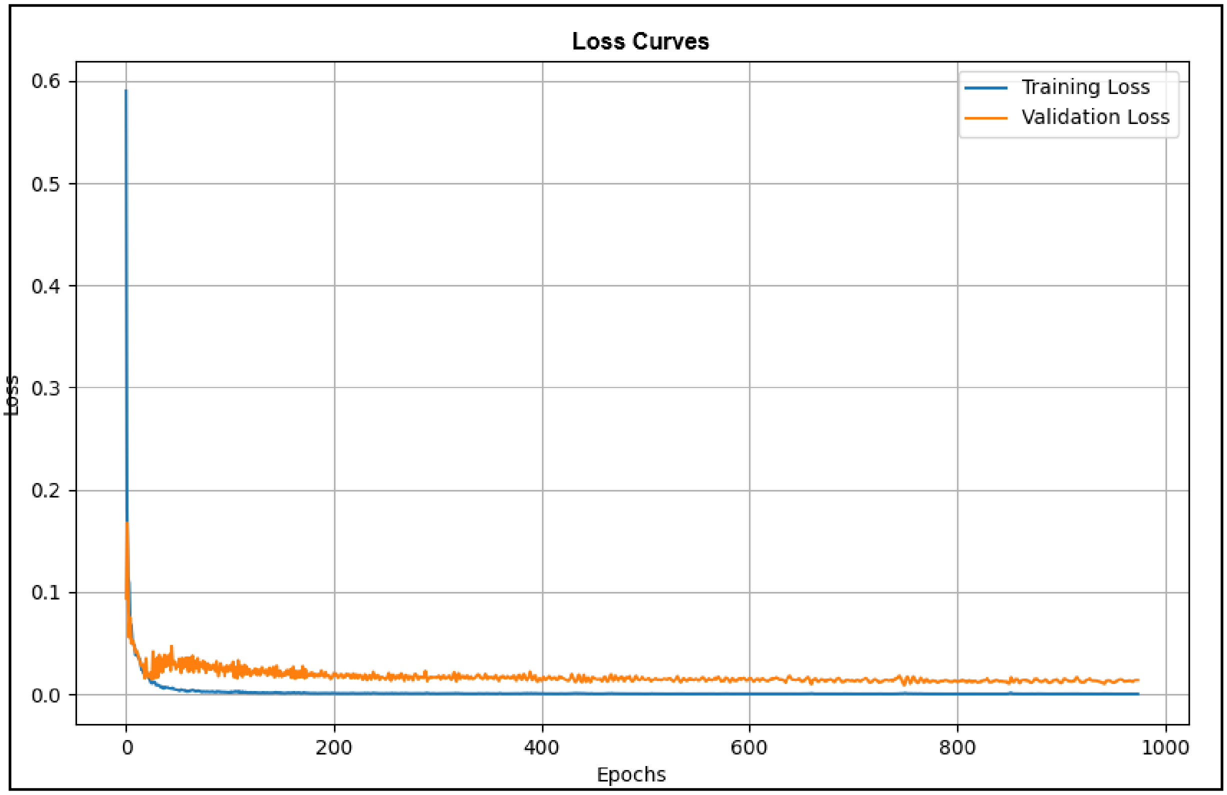

Figure 30: Training Loss and Validation Loss Curves of LSTM Model

The Figure 32 illustrates the prediction error distribution values produced by the LSTM model. Most of the error centers around zero, showing that this model is very accurate and thus exhibits minimal bias in the generated predictions. The majority of errors lie within a narrow range of approximately -0.008 to 0.006, showcasing the model's ability to closely approximate the true values. The slight bimodal nature of the distribution suggests the potential existence of distinct data subgroups, which the model has generally managed to handle well. The small magnitude of the errors and symmetric distribution evidences the robustness of the model and its applicability to this regression task, with only minor deviations evident in a few outlier cases.

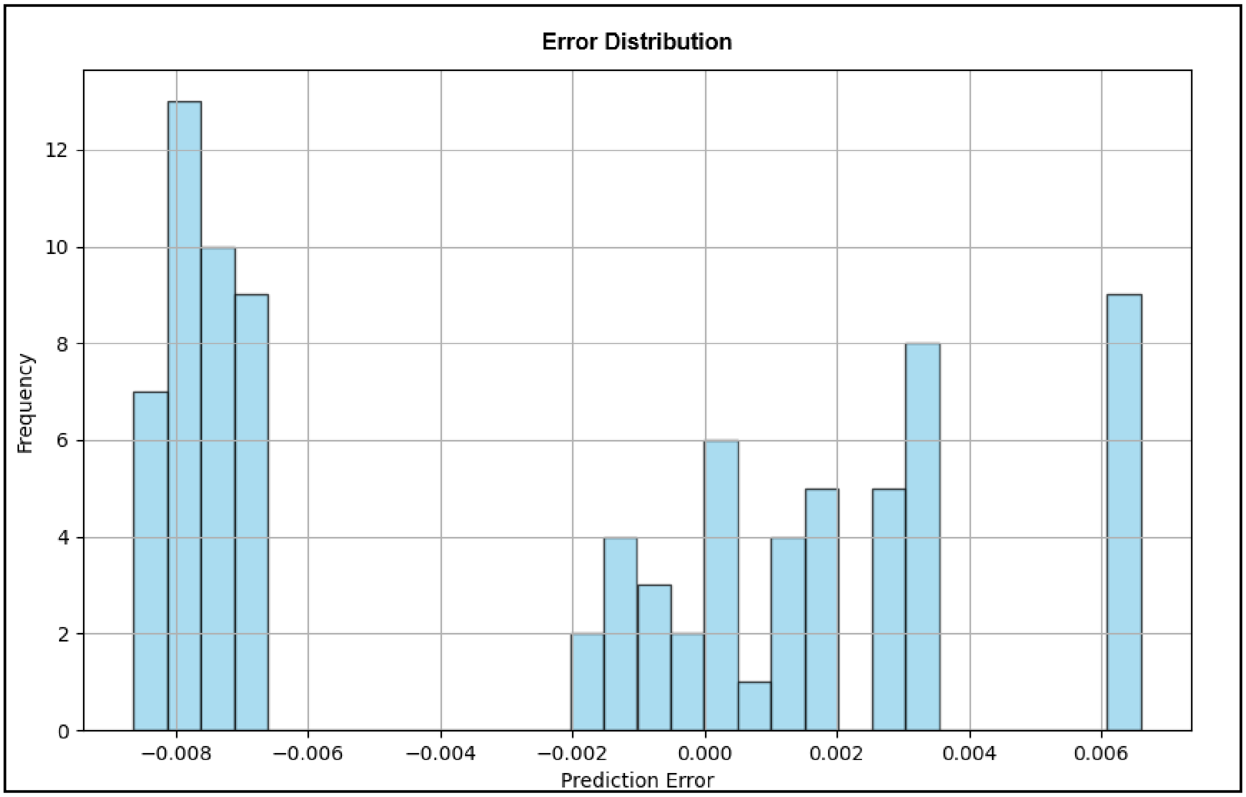

Figure 31: The Prediction Error Distribution values produced by the LSTM model

The LSTM model outperformed all other approaches tested in this Thesis, the LSTM model produced the most accurate and reliable predictions. This is further reflected in its low MAE and RMSE values and its high R^2 score. This is also supported by visualizations that further emphasize the strength of the model for possible deployment in real-world predictive tasks.

5.2 Models Performance Comparison

The standalone models, including RF, KNN, and GBR, served as a foundational benchmark for evaluating performance. Among them, GBR achieved the lowest MSE of 16.76, indicating its superior predictive accuracy. RF followed closely with an MSE of 16.79, thus putting up a performance quite close to that of GBR. On the other hand, KNN performed very poorly, with a significantly higher MSE of 49.11, reflecting its poor ability to model the underlying patterns in the data effectively. Similarly, the R^2 scores reinforce this observation; GBR attained the highest R^2 score of 0.256, slightly outperforming RF at 0.255, while KNN yielded a negative R^2 score of -0.11, meaning that it did not identify any relationship at a worthwhile significance for the data. These results highlight the effectiveness of ensemble-based methods like RF and GBR in achieving better accuracy and robustness in comparison to distance-based methods like KNN.

Table 21: Performance Comparison of RF, KNN, and GBR Model

Model Performance	RF	KNN	GBR
Mean Squared Error	16.79	49.11	16.76
R^2 Score	0.255	-0.11	0.256

Moving to other models, the XGBoost and SVM+RF ensemble classification tasks demonstrated improved performance compared to their counterparts. The SVM + RF model achieved an accuracy of 0.68, significantly outperforming XGBoost, which had an accuracy of 0.4318. Similarly, in terms of the F1 score, which balances precision and recall, the SVM + RF model showed stronger performance with a score of 0.66 compared to XGBoost 0.42. This trend is further reinforced by the precision and recall scores: while the SVM + RF model reached a precision of 0.59 and a recall of 0.45, XGBoost achieved a precision of 0.4434 and a recall of 0.4318. These results indicate that the ensemble-based hybrid approach of SVM + RF effectively leverages the strengths of both algorithms, resulting in more reliable and accurate classification outcomes compared to the standalone XGBoost model. However, these ensemble models still fell short of the more advanced methods tested later in the evaluation process.

Table 22: Performance Comparison of XGBoost and SVM+RF ensemble Model

Model Performance	XGBoost	SVM + RF
Accuracy	0.4318	0.68
F1 Score	0.42	0.66
Recall Score	0.4318	0.45
Precision Score	0.4434	0.59

The performance comparison between the GBR + XGBoost ensemble model and the XGBoost + CatBoost + SVR + LightGBM ensemble model highlights the superior predictive capability. The GBR + XGBoost model achieved a significantly lower MSE of 1.31 compared to 5.32 for the XGBoost + CatBoost + SVR + LightGBM ensemble model, indicating a more accurate fit to the data. Furthermore, for GBR + XGBoost, the RMSE was 1.81 which was slightly lower as compared to that of the XGBoost + CatBoost + SVR + LightGBM ensemble model with RMSE of 1.99, which manifests its ability towards minimizing the prediction errors. In terms of explanatory power, GBR + XGBoost achieved a higher R^2 Score of 0.75, surpassing the R^2 Score of 0.72 for the XGBoost + CatBoost + SVR + LightGBM ensemble. The inclusion of diverse algorithms in this ensemble enhanced its capacity to model various aspects of the data, demonstrating the effectiveness of leveraging heterogeneity in predictive modeling. These results demonstrate that the GBR + XGBoost model offers better overall performance in terms of accuracy and reliability.

Table 23: Performance Comparison of GBR + XGBoost ensemble Model, XGBoost + CatBoost + SVR + LightGBM ensemble Model and LSTM Model

Model Performance	GBR + XGBoost	XGBoost + CatBoost + SVR + LightGBM	LSTM
MSE	1.31	5.32	0.0000325
RMSE	1.81	1.99	0.0057
R^2 Score	0.75	0.72	0.9995

The LSTM model emerged as the best-performing approach, delivering exceptional results with an MSE of 0.000032, an R^2 value of 0.9995, and an RMSE of 0.0057. This model outperformed all the other models by a big margin, topping both in accuracy and minimizing the error rate. Moreover, the LSTM achieved an impressive R^2 Score of 0.9995, signifying that it nearly perfectly explains the variance in the target variable. Its sequential processing capabilities allowed it to effectively capture temporal dependencies and complex nonlinear patterns within the data, making it a superior choice for this task. The loss curve of the LSTM model and the visualization also showed the consistency and precision of actual versus predicted values, making it stand unparalleled in this thesis.

The evaluation of different models showed notable variations in their predictive performance, with the LSTM model standing out as the most effective option. While ensemble-based models such as GBR + XGBoost and XGBoost + CatBoost + SVR + LightGBM displayed competitive accuracy and robustness, LSTM outperformed them all with its remarkable ability to capture temporal dependencies, achieving nearly flawless results. Not only were the evaluation metrics carefully analyzed, but the prediction outputs of all models were also examined. Among these, the LSTM model consistently produced predictions that were almost perfect, reinforcing its reputation as the most dependable and effective model. These results highlight the capabilities of advanced machine learning models, especially LSTM, in delivering both high accuracy and precise predictions, making it the optimal choice for fine-tuning roll grinding machine parameters.

6. Discussion

The focus of this thesis has been to develop the use of advanced predictive analytics to optimize roll grinding processes with respect to minimal form error and simultaneously optimizing grinding time. By employing a variety of machine learning models and ensemble techniques, the thesis aimed to enhance the precision and efficiency of predictions related to roll grading, a critical aspect of industrial operations. The findings highlight how the integration of sophisticated modeling approaches can address longstanding challenges in roll grinding, enabling industries to achieve better outcomes in terms of quality, productivity, and cost-effectiveness.

The grinding process is influenced by several key parameters that together determine its overall efficiency. Among the most important factors identified in this thesis are the *type of machine*, the *grinding wheel* used, and the *steps involved in the grinding process*. The choice of grinding machine directly affects the precision, stability, and speed of the operation, which in turn influences the quality and efficiency of the outcome. Similarly, selecting the right grinding wheel considering its *material*, *size*, and abrasiveness is crucial for optimizing cutting efficiency, heat generation, and *surface finish*. Additionally, the number and order of grinding steps are essential, as they determine both the time needed and the accuracy of the process. An optimized sequence minimizes redundancy, reduces the risk of errors, and helps achieve a consistent and precise result.

Other important parameters include *grinding time*, the *sequence of steps*, and the *type of roll* being processed. The duration of grinding impacts both operational efficiency and resource use, with excessive grinding leading to unnecessary wear and energy consumption. The type of roll, including its material, hardness, and composition, requires specific adjustments to the grinding parameters to ensure the best results. Furthermore, factors such as *surface roughness*, *feed rate*, and the *contact time between the wheel and roll* are critical for achieving the desired surface finish and maintaining effective material removal. Surface roughness is directly influenced by the choice of grinding wheel, feed rate, and overall process settings. The feed rate, in particular, significantly affects the cutting force, heat generation, and quality of the final product. The contact time between the wheel and the roll determines how much material is removed and the risk of overheating or thermal damage to the roll. Finally, the *shape of the roll* adds complexity to the grinding process, where irregularities or customized profiles require careful consideration.

Advanced predictive models implemented in the process have greatly enhanced the accuracy of the predictions, hence a more accurate estimation of roll grading parameters. This improvement enables industries to proactively identify deviations and anomalies, minimizing the occurrence of defects and reducing the need for corrective actions. By optimizing grinding time and reducing form error, the models contribute to enhanced operational efficiency, ensuring consistent quality in roll products while conserving resources. Such advancements reflect the potential of predictive analytics to revolutionize traditional manufacturing processes.

The development of these predictive models involved using advanced methods like ensemble techniques and deep learning models, particularly LSTM networks. These models were trained on large datasets, allowing them to capture the complex relationships between operational parameters and their effects on grinding outcomes.

Employing ensemble techniques such as Random Forest and Gradient Boosting improved model accuracy by integrating multiple base learners, which helped reduce overfitting and enhance generalizability. On the other hand, the LSTM model excelled at capturing temporal dependencies, making it especially effective for predicting optimal grinding parameters in time-series data, thereby minimizing both grinding errors and processing time. These predictive models underwent validation and fine-tuning to deliver accurate and reliable recommendations for the best grinding parameters. Consequently, industries can proactively optimize their grinding processes, leading to fewer form errors, enhanced product quality, and significant time savings in operations. Predictive analytics thus emerges as a crucial tool for achieving optimal grinding performance and streamlining industrial workflows.

The comparative analysis of various machine learning models in this thesis demonstrates the importance of selecting the most appropriate predictive approach based on the specific requirements of the roll grinding process. Contrasting with other simpler models, the application of ensemble techniques with more sophisticated algorithms strengthens their accuracy and robustness by catching a wide range of complex relationships. These findings emphasize the role of advanced analytics in driving innovation and continuous improvement in industrial contexts.

The use of data-driven predictive models in the grinding process has brought about significant advantages, such as better product quality, lower defect rates, and improved operational efficiency. By accurately forecasting the best grinding parameters, these models help reduce form errors, ensuring that the rolls produced align with the required specifications and decreasing the need for expensive rework or corrective measures. Furthermore, the models streamline grinding time, resulting in quicker production cycles and more effective resource use. They also facilitate a transition from reactive to proactive maintenance strategies, predicting equipment requirements and minimizing downtime. In addition to operational enhancements, predictive models have led to tangible improvements in cost-effectiveness by reducing errors, limiting excessive grinding, and optimizing energy and material consumption, which contributes to more sustainable practices. Ultimately, leveraging predictive analytics not only boosts the efficiency and quality of grinding operations but also promotes long-term sustainability and cost savings in manufacturing.

Another important advantage of these predictive analytics methods is that they can enable proactive decision-making and predictive maintenance strategies. By providing reliable forecasts of roll grading outcomes, these models allow for targeted interventions and optimized maintenance schedules. This shift from reactive to proactive operations extends machinery and equipment life, besides reducing downtime and associated costs, thereby directly contributing to improved asset utilization and productivity.

The use of advanced predictive analytics for optimizing roll grinding processes presents a transformative opportunity for the industry. By reducing form error, optimizing grinding time, and enabling proactive maintenance, these methods pave the way for more efficient, sustainable, and high-quality operations. The insights gained from this thesis underscore the potential of predictive modeling to reshape traditional industrial practices, offering a blueprint for future advancements in manufacturing and beyond.

7. Conclusion and Future Work

7.1 Conclusion

This thesis has investigated the use of advanced predictive analytics in the optimization of the roll grinding processes, focusing on reducing form error and optimizing grinding time. Of the various machine learning models and techniques tried, the LSTM network emerged as the most successful among all for achieving perfection in these objectives. This network is incomparable in capturing sequential dependencies with complex temporal patterns in analyzing intricate relationships between the grinding parameters and process outcomes.

The LSTM model's predictions enabled significant improvements in both form error reduction and grinding time optimization. By providing precise recommendations for parameter settings such as feed rate, spindle speed, pressure and depth of cut, the model effectively minimized deviations in form error while maintaining high-quality surface finishes. Simultaneously, optimizing grinding time has made for an improved process efficiency without quality degradation, which still is one of the key challenges in modern manufacturing.

The results of the thesis reveal a transformation potential brought about by data-driven approaches to achieve precision and efficiency in roll grinding. The LSTM model, using high-frequency sensor data and combining it with machine-specific parameters, made the process optimization proactive and systematic. This has not only reduced operational variability but also enhanced productivity and cost efficiency for competitive advantages in precision manufacturing.

Moving forward, this research lays the foundation for further advancements in predictive analytics and process control. Incorporating IoT-driven real-time monitoring and integrating the LSTM model with digital twin technologies can provide even greater insights into process dynamics. These innovations hold the potential to revolutionize roll grinding by enabling real-time adjustments, improving sustainability, and further aligning manufacturing practices with the demands of precision, consistency, and efficiency.

7.2 Future Work

The rapidly transforming field of predictive modeling in production environments is driven by transformational advances in machine learning, availability of data, and computing capabilities. This evolution opens up a host of future research and innovation avenues for refining system performance and capabilities with an eye to overcoming practical challenges in industrial applications. This section considers in-depth various potential directions for the expansion of the existing work, focusing on enhancing system efficiency, using advanced methodologies, and combining different data sources for more powerful and effective solutions.

7.2.1 Exploration of Additional Machine Learning Techniques

For future research leading to enhanced predictability and improved generalization in production environments, the work could be done with the introduction of advanced machine learning techniques. These approaches are better positioned to handle specific challenges peculiar to industrial settings, like non-linear variable relations, sparsity in data, and noisiness. The result can be stronger, more interpretable, and efficient systems for solving sophisticated production scenarios.

Ensemble Learning provides a potent method of improving the robustness of predictions by modeling in a collective strength coming from multiple models. Techniques such as Random Forest, Gradient Boosting, and Stacking operate on the principle of aggregating predictions from several individual models to produce a consensus outcome. By doing so, this aggregation reduces variance and bias, hence yielding better accuracy in predictions. For example, Random Forest constructs a large amount of decision trees and outputs the mode or average of their predictions. hence, it is stable even if noisy data is present. While Gradient Boosting builds the models in an iterative way and corrects the errors of the previous one, turning the algorithm very powerful for tasks with a lot of tricky patterns in data. These ensemble methods are of particular importance in production due to diverse data properties that require flexible yet reliable models.

Unsupervised Learning has been used effectively to provide insight into unlabeled datasets by learning hidden structures and patterns. Clustering techniques, such as k-means or hierarchical clustering, may group data points together based on similarities. hence, anomaly detection could be done by pointing out instances which significantly deviate from typical clusters. Dimensionality reduction methods, including t-SNE and PCA, simplify complex datasets by projecting them onto lower-dimensional spaces. This not only enables visualization but also helps draw out meaningful features for fault detection or predictive maintenance.

Another promising direction for optimizing dynamic processes in production is Reinforcement Learning (RL). Unlike classical machine learning models, their algorithms learn from interacting with the environment by receiving rewards or penalties to direct their actions. This approach is particularly fitted for resource allocation, scheduling, and adaptive quality control, where decisions are made in real time. For instance, RL can dynamically allocate resources to production lines based on workload, minimizing bottlenecks and maximizing throughput. Using continuous learning from the environment, RL systems can respond to changes in production demands and, thus, find a high field of applications in flexible manufacturing.

Future work should focus on adapting these advanced techniques to the specific needs of industrial applications. This involves algorithm modifications to deal with the scale and complexity of production data, the integration of domain knowledge to enhance model relevance, and ensuring computational efficiency to realize real-time or near-real-time predictions. Coupled together, these methodologies think of ensemble learning coupled with explain ability tools or reinforcement learning working with unsupervised clustering could actually unlock new frontiers in making systems smarter and adaptable. Addressing avenues such as improving reliability, efficiency, and the impact of production environments allows predictive modeling to evolve.

7.2.2 Advancing Predictive Modeling in Modern Production Systems: Real-Time Insights, Integration, and Innovation

Predictive modeling is experiencing an epoch transformation in the production environment propelled by ever-capable machine learning, growing data availability, and high-performance computing technologies, unparalleled opportunities thereby created for upgrading systems to higher scalability, flexibility, and performance become readily possible. In fact, key discussion will start over setting up the real-time predictive system that in real time preprocess data for anomaly identification, dynamic optimization of resources, and reduction in downtime. Latency, speed of algorithms, and accuracy-such challenges can be overcome with the help of state-of-the-art technologies: edge computing, cloud platforms, and stream processing frameworks such as Apache Kafka or Flink. Examples of transformational applications that could come out of these systems are real-time predictive maintenance and dynamic adjustment of workflows for high operational efficiency and resource optimization.

Besides, integrating the predictive models in wider production pipelines unlocks synergies that enable seamless coordination of supply chain logistics, quality assurance, and sustainability practices. Digital twins and similar technologies provide a virtual framework for integrating data flows, process simulation, and holistic optimization. Transfer learning further extends the adaptability of predictive systems by enabling models to generalize across domains, hence rendering them of particular value to small-scale manufacturers with limited data. Moreover, the integration of diverse data sources on environmental conditions, operators' feedback, and market demand will enrich the predictive capabilities. Although the issue of heterogeneity still partially exists, advances in the techniques of data fusion promise comprehensive and insightful analyses. These progressive directions herald a future where predictive modeling underpins smarter, more efficient production systems, meeting the evolving demands of modern industry with precision and resilience.

8. Bibliography

[1] Liang, X., Xu, X., Wang, X. "Advancements in Machine Tool Technology for High-Precision Manufacturing." *Journal of Manufacturing Processes*, vol. 58, pp. 457-468, 2022. doi: 10.1016/j.jmapro.2021.11.014.

[2] Brettel, M., Friederichsen, N., Keller, M., Rosenberg, M. "How Virtualization, Decentralization, and Network Building Change the Manufacturing Landscape: An Industry 4.0 Perspective." *International Journal of Information and Communication Engineering*, vol. 8, no. 1, pp.37-44, 2014. [Online]. Available: https://www.researchgate.net/publication/285495324_How_virtualization_decentralization_and_network_building_change_the_manufacturing_landscape_An_Industry_40_Perspective

[3] Yin, Z., Xu, X., Wang, L. "Energy-Efficient Machine Tools in Industry 4.0: Framework and Methodologies." *Journal of Manufacturing Systems*, vol. 51, pp. 287-299, 2019. doi: 10.1016/j.jmsy.2019.06.001.

[4] Koren, Y., Shpitalni, M. "Design of Reconfigurable Manufacturing Systems." *Journal of Manufacturing Systems*, vol. 29, no. 4, pp. 130-141, 2010. doi: 10.1016/j.jmsy.2011.01.001.

[5] Haraguchi, N., Cheng, C. F. C., Smeets, E. "The Importance of Manufacturing in Economic Development: Has This Changed?" *World Development*, vol. 93, pp. 293-315, 2017. doi: 10.1016/j.worlddev.2016.12.013.

[6] Klocke, F., Brecher, C., Schröer, U. "Innovative Machine Tool Design for Enhanced Productivity and Sustainability." *CIRP Annals*, vol. 67, no. 2, pp. 705-728, 2018. doi: 10.1016/j.cirp.2018.05.001.

[7] Wang, L., Li, S. "Green Machining: Sustainable Manufacturing Practices in the Context of Machine Tool Production." *Journal of Cleaner Production*, vol. 231, pp. 1-15, 2019. doi: 10.1016/j.jclepro.2019.05.178.

[8] Kumar, P., Kaushik, S. "Predictive Maintenance and Data-Driven Process Optimization in Manufacturing." *International Journal of Advanced Manufacturing Technology*, vol. 103, no. 9, pp. 4167-4183, 2019. doi: 10.1007/s00170-019-03940-x.

[9] Zhao, Y., Chen, J. "IoT-Based Remote Monitoring and Control Systems in Smart Manufacturing." *Proceedings of the 2019 International Conference on Manufacturing Systems*, 2019, pp. 234-239. doi: 10.1109/ICMS.2019.00119.

[10] Li, Y., Wang, L. "Application of Production Data Analysis in Improving Machine Tool Production Efficiency." *Proceedings of the 2020 IEEE International Conference on Industrial Technology*, 2020, pp. 1730-1735. doi: 10.1109/ICIT.2020.7437759.

[11] Takyar, A., Takyar, A. "*How to build a predictive ML model for manufacturing operations?"* LeewayHertz - AI Development Company. (2023, May 13). [Online]. Available: https://www.leewayhertz.com/how-to-build-predictive-ml-model-for-manufacturing/

[12] Gupta, S., Soni, P. "Production Planning and Scheduling: Enhancing Quality Control and Resource Utilization." *Proceedings of the 2017 IEEE International Conference on Industrial Engineering and Engineering Management*, 2017, pp. 1184-1188. doi: 10.1109/IEEM.2017.8292372.

[13] *1 Industry 4.0 CMMS for integration. Get One View | MaintWiz*. (n.d.). maintwiz.com. [Online]. Available: https://www.maintwiz.com/connected-systems/

[14] Xia, Q., Tang, Y. "Improving Manufacturing Efficiency and Sustainability through Advanced Analytics and Smart Technologies." *Proceedings of the 2019 IEEE International Conference on Advanced Manufacturing*, pp. 105-113. doi: 10.1109/ICAM.2019.00123.

[15] Li, X., Liu, Y. *"Reducing Downtime and Lowering Production Costs in Manufacturing Systems: A Data-Driven Approach." International Journal of Production Economics*, vol. 217, pp. 59-72, 2019. doi: 10.1016/j.ijpe.2019.05.006.

[16] Ronen, D., Chen, H. *"Data-Driven Approaches to Improving Grinding Performance and Reducing Cycle Time." Journal of Manufacturing Processes*, vol. 34, pp. 410-421, 2018. doi: 10.1016/j.jmapro.2018.06.005.

[17] Gao, W., Li, Z**.** *"Data-Driven Optimization of Grinding Parameters: A Predictive Model-Based Approach." International Journal of Advanced Manufacturing Technology*, vol. 106, no. 9-12, pp. 3849-3860, 2020. doi: 10.1007/s00170-020-05394-3.

[18] Smith, J., Brown, R. *"Data-Driven Innovations in Roll Grinding Operations: An Industrial Perspective." Journal of Manufacturing Science and Engineering*, vol. 142, no. 3, pp. 1-12, 2020. doi: 10.1115/1.4046821.

[19] Liu, X., Zhang, W. *"Integration of Machine Sensors and Drives for Real-Time Data Collection in Grinding Machines." Journal of Intelligent Manufacturing*, vol. 32, no. 6, pp. 1759-1773, 2021. doi: 10.1007/s10845-021-01747-8.

[20] Brecher, C., Esser, M. *"Machine Tool Communication and Data Integration with EtherCAT in Grinding Applications." Journal of Advanced Manufacturing Systems*, vol. 16, no. 1, pp. 45-62, 2018. doi: 10.1142/S0219686718500045.

[21] Müller, R., Schuler, P. *"Data Communication in Automated Machine Tools: ADS Protocol and Its Applications." International Journal of Automation Technology*, vol. 14, no. 3, pp. 198-210, 2019. doi: 10.20965/ijat.2019.p0198.

[22] Klocke, F., Kuchle, A., Breuer, D. *"Roll Grinding with Advanced Data-Driven Techniques: Applications in Predictive Maintenance." International Journal of Advanced Manufacturing Technology*, vol. 78, no. 3-4, pp. 691-704, 2015. doi: 10.1007/s00170-015-6720-1.

[23] Brecher, C., Esser, M., Witt, S. *"Data-Driven Optimization in Roll Grinding: Applications of Machine Learning Algorithms." Production Engineering*, vol. 4, no. 3, pp. 287-294, 2010. doi: 10.1007/s11740-010-0230-y.

[24] Zhou, J., Sun, G. *"Production Data Collection and Utilization in Roll Grinding Operations." Journal of Manufacturing Processes*, vol. 41, pp. 123-130, 2019. doi: 10.1016/j.jmapro.2019.01.010.

[25] Klocke, F., Kuchle, A., Breuer, D. *"Advances in Machine Monitoring for Precision Grinding." International Journal of Advanced Manufacturing Technology*, vol. 81, pp. 689-702, 2015. doi: 10.1007/s00170-015-7239-1.

[26] Klocke, F., Kuchle, A., Breuer, D. *"Analyzing Performance Data in Precision Roll Grinding." International Journal of Advanced Manufacturing Technology*, vol. 83, pp. 801-812, 2016. doi: 10.1007/s00170-016-7513-2.

[27] Brecher, C., Merz, M., Fey, M. "Tool Wear Monitoring in Roll Grinding Operations: Enhancing Precision and Tool Life." CIRP Annals - Manufacturing Technology**,** vol. 67, no. 1, pp. 317-320, 2018. doi: 10.1016/j.cirp.2018.04.062.

[28] Zhou, J., Li, X. "Data Management for Maintenance and Condition Monitoring in Roll Grinding Machines." Journal of Manufacturing Systems**,** vol. 45, pp. 87-98, 2018. doi: 10.1016/j.jmsy.2018.01.002.

[29] Kumar, S., Singh, A. P. "Vibration Analysis for Predictive Maintenance in Grinding Machines." Journal of Vibration and Control**,** vol. 24, no. 2, pp. 105-120, 2018. doi: 10.1177/1077546317720771.

[30] Ramesh, R., Mannan, M. A., Poo, A. N. "Error Compensation in Machine Tools — A Review: Part II: Thermal Errors." International Journal of Machine Tools and Manufacture**,** vol. 40, no. 9, pp. 1257–1284, 2000. doi: 10.1016/S0890-6955(99)00122-2.

[31] Amini, S., Ahmadi, M. "Energy Efficiency in Grinding: Theoretical and Practical Insights." *Journal of Manufacturing Processes*, vol. 56, pp. 120-130, 2020. doi: 10.1016/j.jmapro.2020.04.006.

[32] Zhou, Y., Zhang, J. "Smart roll grinding and data-driven decision support in advanced manufacturing." *Journal of Manufacturing Processes*, vol. 56, pp. 130-142. (2020). doi: 10.1016/j.jmapro.2020.03.004.

[33] M. S. S. M. Rahman, "Development and Optimization of Grinding Technology in the Manufacturing Industry," *International Journal of Advanced Manufacturing Technology*, vol. 56, no. 5–8, pp. 675-690, 2011. doi: 10.1007/s00170-010-2689-4.

[34] Zhou, J., Sun, G. "Key Performance Indicators in Roll Grinding: A Data-Centric Approach." *Journal of Manufacturing Processes, 46*, 121-133. (2020). doi: 10.1016/j.jmapro.2020.02.003

[35] Attaran, M., Attaran, S. "Opportunities and challenges of implementing predictive analytics for competitive advantage". *International Journal of Business Intelligence Research*, *9*(1), 1-26. (2018). https://doi.org/10.4018/IJBIR.2018070101

[36] James, G., Witten, D., Hastie, T., Tibshirani, R. "*An Introduction to Statistical Learning: With Applications in R"*. Springer. https://doi.org/10.1007/978-1-4614-7138-7

[37] Ashraf, R., Hussain, M., Jaffar, A., Ramay, W., Faheem, M. (2024). "Revolutionizing Urdu Sentiment Analysis: Harnessing the Power of XLM-R and GPT-2". *IEEE Access*, PP, 1-1. (2013). https://doi.org/10.1109/ACCESS.2024.3429496

[38] Zhu, X., Goldberg, A. B. "Introduction to Semi-Supervised Learning," *IEEE Transactions on Knowledge and Data Engineering*, 21(9), 1353-1366. (2009). https://doi.org/10.1109/TKDE.2009.11

[39] Hochreiter, S., Schmidhuber, J. "*Long Short-Term Memory"*. Neural Computation, 9(8), 1735-1780. (1997). https://doi.org/10.1162/neco.1997.9.8.1735

[40] Improving Convection Trigger Functions in Deep Convective Parameterization Schemes Using Machine Learning. (n.d.). *Scientific Figure*. ResearchGate. [Accessed 21 Jan 2025]. [Online]. Available: https://www.researchgate.net/figure/Flow-diagram-of-gradient-boosting-machine-learning-method-The-ensemble-classifiers_fig1_351542039

[41] Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y. *Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation*. In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2014), 1724-1734. [Online]. Available: https://www.aclweb.org/anthology/D14-1179

[42] Riebesell, J. *Random Forest*. (2022, April 10). [Online]. Available: https://tikz.net/random-forest/

[43] S. R, "Understand Random Forest Algorithms with Examples," Jun. 17, 2021. [Online]. Available: https://www.analyticsvidhya.com/blog/2021/06/understanding-randomforest/#:~:text=A.%20Random%20Forest%20is%20a,model%20is%20a%20decision%2 0tree.

[44] Boldini, D., Grisoni, F., Kuhn, D., Friedrich, L., Sieber, S. "Practical guidelines for the use of gradient boosting for molecular property prediction". *Journal of Cheminformatics, 15*, Article 7. (2023). https://doi.org/10.1186/s13321-023-00743-7

[45] Chen, T., Guestrin, C. "*XGBoost: A scalable tree boosting system"*. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785-794. (2016). https://doi.org/10.1145/2939672.2939785

[46] Ke, G., Meng, Q., Finley, T., Wang, T., Yu, D. "*LightGBM: A highly efficient gradient boosting decision tree"*. Advances in Neural Information Processing Systems, 30. (2017). [Online]. Available: https://arxiv.org/abs/1712.09913

[47] Prokhorenkova, L. A., Gusev, G., Vorobyev, M., et al. "*CatBoost: unbiased boosting with categorical features"*. Proceedings of the 32nd International Conference on Neural Information Processing Systems, 6638-6648. (2018). https://doi.org/10.5555/3327757.3327846

[48] Zhao, F. "A Literature Study of Asset Pricing Based on Machine Learning Method". *Advances in Economics, Management and Political Sciences*, 34, 27-36. (2023). https://doi.org/10.54254/2754-1169/34/20231669

[49] Khan, M., Islam, S., Sarkar, S., Ayaz, F., Kabir, M., Albraikan, A., Almalki, F. "Machine Learning Based Comparative Analysis for Breast Cancer Prediction". *Journal of Healthcare Engineering*, 2022, 1-15. https://doi.org/10.1155/2022/4365855

[50] Kumar, A., Sharma, R. "Predicting roll grinding quality using machine learning techniques: A comparative study of Support Vector Machines (SVM) and K-Nearest Neighbors (KNN)". *Journal of Manufacturing Processes*, 58, 1-10. (2020). https://doi.org/10.1016/j.jmapro.2020.03.010

[51] Guyon, I., Elisseeff, A. "*An Introduction to Variable and Feature Selection". Journal of Machine Learning Research, 3*, 1157-1182. (2003). [Online]. Available: https://www.jmlr.org/papers/volume3/guyon03a/guyon03a.pdf

[52] Jain, A., Nagpal, A. "*A Comprehensive Survey of Feature Scaling and Transformation Techniques for Machine Learning". Journal of Computer Science and Technology, 35*(1), 45-60 (2020).

[53] Jolliffe, I. T. "*Principal Component Analysis* (2nd ed.)". (2002). Springer Series in Statistics. Springer. doi**:** 10.1007/b98835

[54] *"Practical guide to the confusion matrix for AI*. (n.d.)". [Online]. Available: https://en.innovatiana.com/post/understand-confusion-matrix-in-ai

[55] Zia, S., Singh, A., Ul-Haq, A., Raj, D. S., "Data-driven approaches for predictive maintenance in manufacturing: A survey". *International Journal of Advanced Manufacturing Technology*, 112(1), 35-53. et al. (2021). https://doi.org/10.1007/s00170-020-06160-7

[56] Sawant, C. "Understanding Feature Scaling in Machine Learning: Standardization vs Normalization". (2023, September 14). *Medium*. [Online]. Available: https://medium.com/@chaitanyasawant/understanding-feature-scaling-in-machine-learning-standardization-vs-normalization-800ad25b0c24

[57] Bergstra, J., Bengio, Y. "Random Search for Hyper-Parameter Optimization. *Journal of Machine Learning Research"*, 13(1), 281-305. (2012). [Online]. Available: https://www.jmlr.org/papers/volume13/bergstra12a/bergstra12a.pdf

[58] *"Strategies to Enhance ML Model Accuracy: A Guide*. (n.d.)". [Online]. Available: https://www.markovml.com/blog/model-accuracy

[59] FasterCapital. (n.d.). *What are the benefits of using machine learning in business?* [Online]. Available: https://fastercapital.com/topics/what-are-the-benefits-of-using-machine-learning-in-business.html

[60] Omarzai, Fraidoon. “Principal Component Analysis (PCA) in Depth - Fraidoon Omarzai - Medium.” *Medium*, 20 Nov. 2024, [Online]. Available: medium.com/@fraidoonomarzai99/principal-component-analysis-pca-in-depth-93c871f25dfa.

[61] Herkules Group, “*CNC Roll Grinding Machine*” [Online]. Available: www.herkules-machinetools.com.

[62] Chavan, Rahul. “Understanding Train, Test, and Validation Split in Simple Quick Terms.” *Medium*, 25 June 2023, [Online]. Available: medium.com/@rahulchavan4894/understanding-train-test-and-validation-dataset-split-in-simple-quick-terms-5a8630fe58c8.