Optimizing Production Efficiency in Manufacturing using Big Data and AI/ML

[1]Venkata Nagesh Boddapati, [2]Sanjay Ramdas Bauskar, [3]Chandrakanth Rao Madhavaram, [4]Eswar Prasad Galla, [5]Janardhana Rao Sunkara, [6]Hemanth Kumar Gollangi

[1]Microsoft Sr. Technical Support Engineer.
[2]Pharmavite LLC Sr. Database Administrator.
[3]Microsoft Support Escalation Engineer.
[4]Depatment of Comp. Sci. University of Central Missouri.
[5]Siri Info Sol. Inc. Sr. Oracle DB Admin.
[6]KPMG Consultant.

[1]venkatanageshboddapati@yahoo.com, [2]sanjayramdasbauskar@outlook.com, [3]Craoma101@outlook.com, [4]EswarPrasadGalla@outlook.com, [5]JanardhanaRaoSunkara@outlook.com, [6]HemanthKumarGollangi12@outlook.com

Abstract

Manufacturers need to ensure that their facilities are in optimal working condition to achieve maximum output. The ideal scenario is to have no wasted time, no mistakes, and a consistently perfect product. Yet, we know that inefficiencies may occur that can put a stop to production, lower profits, and result in substandard outputs. Market trends and dynamics can easily drive the relevant factors that allow for a product's production efficiency to change. What if new regulations impact material purity and lead to significantly more downtime for cleaning? How will the production volume respond? What drives the number of tool changes? Should we be installing additional machines, or can we change production to mitigate the aforementioned regulation? All of these are questions that can easily be analyzed by using historical data. However, traditional analysis methods can take days or even weeks just to create a production data report before further analysis can take place, so such analysis is often avoided or not as thorough as it could be.

Keywords: Production Efficiency, Manufacturing Optimization, Big Data, AI/ML (Artificial Intelligence/Machine Learning), Operational Downtime, Regulatory Impact, Tool Changes, Historical Data Analysis, Production Volume, Data-Driven Insights.

1. Introduction

Traditionally, production landlords in the manufacturing industry have been highly experienced engineers who have spent years on the shop floor keeping plant processes running optimally during unscheduled downtime. They relied on the help of slightly less experienced engineers who spent a lot of time manually collecting process data and analyzing it. Both sets of engineers built up considerable experiential knowledge on what was likely to go wrong in, say, a rolling mill, forging process, or heat treatment of steel, and took corrective steps as necessary. What was needed was a way to capture all this experiential knowledge and integrate it with live process data in one place so that necessary action could be taken quickly and efficiently. For many years, this was not possible. Over the past 10 to 20 years, we have started to first automate the capturing of process data and now integrate process data with engineering rules and experiential knowledge to come up with intelligent decision-making systems. Today, these systems are based on data analysis tools and advanced computer science. This paper demonstrates why such support is so urgently required.

The usual problems discussed in an advanced industry describe issues associated with the design or anticipated performance of user-defined plants or machinery. However, in a mineral processing or manufacturing environment, the machinery or equipment is highly customized, and the quality of the product is heavily influenced by feed quality, weather conditions, and machine health. It is virtually impossible to control the quality of these variable parameters and very difficult to predict how these quality parameters will affect the maintenance and response to challenges of the final product. Every good-quality IO pellet plant has to prevent a massive losing streak due to using feed ore, which causes breakage of pellets during their handling. Once a massive losing streak has happened and breakage becomes the key issue, it is already good enough to achieve significant amounts of pellets into road sweepings, etc. This paper argues for the use of offline and online data analysis tools for knowledge capture, as well as the application of advanced computer science tools for capturing experiential knowledge as an intelligent dynamic decision support system in the iron ore pellet manufacturing unit. We consider the benefits of the inclusion and expression of expert knowledge in the form of linguistic rules and decision trees in a data analysis tool to be the speed and expression as if an experienced knowledge principal discussed with his or her process trainee calmly and analytically, and the clarity of the answer to "how would you make the treatment of data more advanced?" optional application to non-synchronization, out-of-pattern detection, quality implication, system health, and applicability under conditions of uncertainty. Data analysis tools may also require an audit where probability counts the take of the input by end-user questions and the relation of the input to the output of the system. In the next section, we discuss big data analysis, advanced computer science tools, and previous work in our area of interest. In section 3, we discuss the case study that we have done. In section 4, we consider experimental results. Finally, we draw our conclusions and list points for future work.

1.1 Background and Rationale

In the world of manufacturing, terms like Industry 4.0, automated and smart factories, industrial data analytics, and data-driven approaches have acquired high popularity and have emerged as key focus areas for business leaders. A fire-and-forget approach to manufacturing to meet market demands is the fundamental force acting in this direction, driven by dynamic customer tastes. On a production floor, whenever a machine or equipment goes down due to some fault in one of its components, it has a significant impact on production efficiency, which is measured as the ratio of quantities produced to the time available. The puzzling questions that commonly occur to the managerial team are: 'How can AI/ML and big data help?' and 'Are our systems in place enough to handle manufacturing fault occurrences?' These questions can be addressed to a considerable extent with a combination of different machine learning models and methodologies like predictive maintenance. There is a need for preparing maintenance data, creating models to predict which component is likely to fail, connecting the models to equipment for decision-making, and taking preventive action to ensure that production is not halted. After identifying important parameters and generating data, data-driven modeling can be applied effectively to identify patterns of stress on particular components of equipment to find trends, clusters, and deviations from regular conditions. Benefits of building manufacturing resilience through AI and ML is shown in figure 1.

Figure 1. Benefits of Building Manufacturing Resilience Through AI and ML

1.2 Research Aim and Objectives

A data-driven operational model was developed early on, which, when looked at retrospectively, comprises the basics of data collection, feature induction, devising an operational model with mixed AI/ML and heuristic components, model testing, deployment, and solution proposal, followed by feedback-based continuous improvement. This research aims to build upon this pre-existing model and validate the established knowledge with the benefit of newly emerged possibilities. As it applies to a specific case of cost optimization, it becomes the objective of this paper. This paper presents novel practices, namely asset utilization and maintenance gap cost, as well as asset utilization-maintenance gap cost, to close the most important gaps in traditional OEE calculations. The asset utilization-maintenance gap costs suggest that a company's production system would see upstream benefits when considering asset utilization and maintenance costs more holistically relative to OEEs.

1.3 Scope and Significance

An exhaustive study posited that "big data is the next frontier for innovation, competition, and productivity." The predecessors had estimated nine impactful strategies for large-scale recession management during the late 1990s and early 2000s, which led to explicit and worthwhile analytical activities using advances in information and digital technology. As consumers, humans produce a substantial amount of data daily. The new data production activities require no more than beings constantly being born who automatically exhibit typical features of standard data-producing information capture devices. Population growth, affluence, and the growth of AI/ML abet this enormous data creation, accumulation, processing, and analysis.

Not only is big data efficient in describing socio-economic activities and priority project characteristics, but for most developing countries, it presents a cost-efficient way for compiling population censuses and monitoring projects, as these exercises consume substantial financial and manpower resources. In today's 'datapolis', largely financial, procurement, and construction companies already employ one AI/ML analyst per company. Because of ongoing increased data proliferation, they institute new data governance rules, seek possibilities of reclaiming personal land ownership rights of AI/ML created communities in data cloud estates, redefine organizations' big data focus, and reskill the workforce to benefit from AI/ML functionalities which, somehow, currently still do not provide a broad functional capability in terms of emulating the human thought process. These companies are exploring big data as the very cornerstone in understanding the intrinsic features of populations as agents of change behaviors in their society. The rise of big data represents a transformative opportunity for innovation and productivity, particularly in developing countries where traditional methods of data collection, like population censuses, can be prohibitively expensive and resource-intensive. As individuals generate vast amounts of data daily, driven by factors such as population growth, increased affluence, and advancements in AI and machine learning, organizations are strategically leveraging this data to gain insights into socio-economic dynamics and project characteristics. In sectors like finance, procurement, and construction, the integration of AI/ML analysts has become commonplace, prompting companies to establish robust data governance frameworks and explore new avenues for land ownership rights in digital contexts. Furthermore, as these organizations redefine their big data strategies, they are reskilling their workforces to harness the potential of AI/ML, striving to understand populations not just as data points but as active agents of change within their communities. Despite the challenges in fully replicating human cognitive processes, the ongoing exploration of big data remains pivotal in shaping informed decision-making and fostering societal advancements.

Equation 1-3 shows the Struggling with Control Chart Limits.

Upper Control Limit (UCL):

$$UCL = \bar{p} + \sqrt[3]{\frac{\bar{p}(1-\bar{p})}{n_i}} \tag{1}$$

Center Line (CL):

$$CL = \bar{p} = \frac{\sum p_i}{\sum n_i} \tag{2}$$

Lower Control Limit (LCL):

$$LCL = \bar{p} - \sqrt[3]{\frac{\bar{p}(1-\bar{p})}{n_i}} \tag{3}$$

Where:

- $\bar{p}$ is the average proportion of defects.
- p_i is the proportion of defects in each sample.
- n_i is the sample size for each subgroup.
- UCL and LCL are the upper and lower control limits, respectively, for the p-chart.

2. Literature Review

The research on optimization of production efficiency starts with understanding the data stream in production for drawing inferences from them for risk and asset management. The physics overlaid on data analytics is a co-discipline. Kaizen encompasses defective patterns and can play a role in the design of big data systems. Competitive factors in production and the role of finance in CAPA are generic. R&D, supply chain, manufacturing, and service functions of any organization influence each other. Big data can be harnessed to simulate various relationships to find alignments. Production risk due to Pareto output and constraints on capacity, spending, and time interactions can limit production. Machine faults, downtime, productivity losses, maintenance, spares, methods, and equipment design are interdependent. Data needs to be available for analysis of these linkages. Preventive maintenance and optimization of equipment reliability by strengthening weak areas will lead to higher production efficiency. Tools are needed to differentiate early signs of an imminent equipment failure. They flag a downtime event coupled with high urgency. The high urgency downtimes are especially troubling from the perspective of lost production, whereas the actual machine failures, while able to produce high urgency downtimes, don't always have time to worsen sufficiently, thus producing a machine failure.

2.1 Big Data in Manufacturing

There are several examples showing how big data helps in manufacturing processes. Even if each case is unique, both horizontal and manufacturing analytics cases exist, such as reducing machine downtime, reducing product defects, predicting equipment failures, reducing energy usage, improving quality, and analyzing semiconductors and sensors to improve quality and increase yield performance. CAM site integrations in a single enterprise data warehouse increase the analysis of form factor measurement, enhance equipment efficiency in the foundry, and enable more complex analytics in the 3D NAND production process. Analytics for performing cloud-platform deployment and daily monitoring, as well as alarms for the strategic sulfur room of smart manufacturing operations, are also important. Building smart test solutions, temperature stabilization inside specific tools used in microelectronics, batch processing analysis, integrated fleet management execution, and production performance reporting are key components.

Among the identified opportunities, there was a need to have easy-to-read and critical parameters faster, such as site performance, FDC parameters, and integrated builds. Extensive use of data output reports, quality sampling reports, and reports on incoming products, as well as additional reports that map product generation and utilization, are essential. Complex interaction analytics are implemented in new products and provide product support. Product scheduling reflects the latest capacity data and commercial data to improve business analytics, including inventory with a live information base and logistics between factories and supply chains. A very interesting case is the "work from home" initiative that enables engineers to continue participating in real-time engineering meetings with data accuracy. The results of the use cases are of great value as they demonstrate the successful use of big data in the semiconductor and electronics context. Additionally, deploying and integrating big data analytics into manufacturing, maintaining regular interactions between our company staff and selected suppliers, along with a common program assessment, require a big data architecture strategy to partner on main aspects for deployment and develop a cost-effective solution with a short time to market.Big data is revolutionizing manufacturing processes, particularly in the semiconductor and electronics sectors, by enabling enhanced analytics and operational efficiencies. Key applications include minimizing machine downtime, predicting equipment failures, and optimizing energy consumption, all of which contribute to improved product quality and yield. Integrating data across manufacturing sites into a centralized enterprise data warehouse facilitates complex analytics, such as monitoring form factor measurements and equipment efficiency in foundries. Additionally, smart manufacturing operations benefit from real-time monitoring and alarm systems, particularly in critical areas like sulfur management. The implementation of user-friendly dashboards for key performance indicators, alongside detailed reporting on product quality and logistics, empowers decision-makers with actionable insights. Notably, the "work from home" initiative illustrates the adaptability of engineering teams to collaborate effectively, leveraging accurate data for real-time decision-making. Ultimately, a robust big data architecture strategy is essential for maintaining strong supplier relationships, optimizing production scheduling, and ensuring a swift and cost-effective deployment of analytics solutions across the manufacturing landscape. Figure 2 depicts the Big Data in Manufacturing.

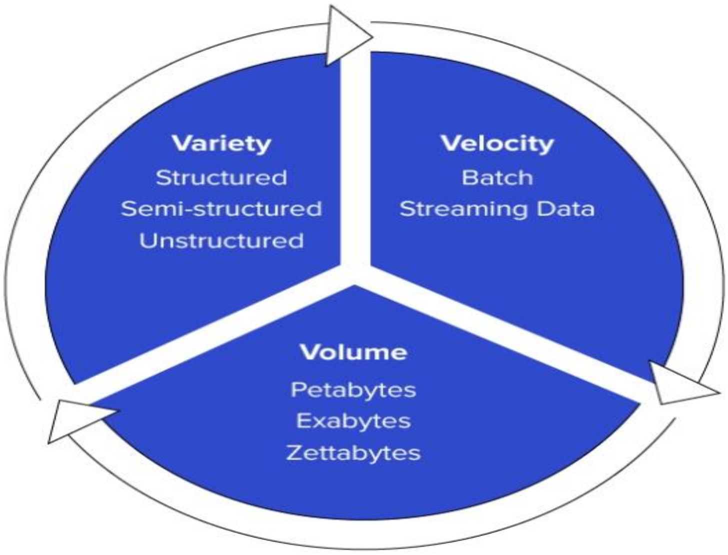

Figure 2. Big Data in Manufacturing

2.2 AI/ML Applications in Production Optimization

Artificial intelligence techniques based on either expert systems, or reinforcement learning alone, or a combination of expert systems and reinforcement learning can be applied in production optimization problems in the manufacturing sector. Here are some examples of these applications: integrating an intelligent decision support system into the production planning process. The purpose of this study is to show how an intelligent system based on an expert system and a reinforcement learning agent can be used to integrate production planning activities. When there are enough computational resources available, multi-agent systems can perform more complex intelligent operations. This can lead to the development of systems for real-time support for online adaptive decision-making that uses intelligent software.

Managing complex robotic manufacturing systems. The implementation of AI technology is performed from the automatic code for generating dispatch rules, as well as the reinforcement learning approach aimed at setting the service conditions for cyclic production and executing production capacity allocations based on the material flow. Automatic generation of the acceptance and dispatch rules of these tasks, with further consideration of reinforcement learning, is performed to optimize the quality and the time of the robotic assembly. Overall, the presented methods of AI integration are highly demanded in industrial sectors such as robotics, where process automation is both necessary and essential for effective performance. The benefits of the intelligent system enhancement include an increase in the quality of the work performed and a decrease in the time of robotic assembling. The structure of the intelligent system planned is designed to be applied to production planning for small batch processes in the field of the mobile industry. It is also capable of renewing production task schedules each time changes in orders and resource allocations occur.

2.3 Integration of Big Data and AI/ML in Manufacturing

In the new era of industry and production, it is common for big data and problems to exceed human expectations and traditional computing modes. People are starting to realize that technology will completely change data acquisition, model establishment, optimization theory, methods, and the mode of interaction; i.e., AI is about the evolution of the four interaction modes among people, the environment, the process, and data. It is necessary to deepen the understanding and applications of production lifecycle data at various levels, with AI being applied in the production management process, and the realization of semantic perceptions of production is expected. It is urgent to speed up the construction of production AI factories. In summary, the application of AI/ML technology in manufacturing promotes horizontal integration rather than solely manufacturers' internal improvement, which significantly reduces conventional time and budget costs associated with process optimization and new product generation. In the future, with the development of the theory and method, optimization efforts will gradually be extended to the lower levels of the manufacturing process. While optimized production will maximize the output of modern manufacturing, its specific goals will evolve from simply ensuring the quantity of product production to ensuring product quality and characteristics. These goals reflect the intelligent and customer-oriented nature of future manufacturing. Additionally, AI technology will help partially ease the mounting pressure on emerging disciplines, such as robotics and blockchain. As AI becomes the leading discipline in computer science, its steps in the field of manufacturing will also become one of the pillars of the Industry 4.0 era.

Equation 4-7 shows the inventory models for certain demand: economic order quantity.

Given values:
H = 12 (holding cost per unit), S = 150 (ordering cost), D = 250,000 (annual demand)

Economic Order Quantity (EOQ):

$$EOQ = Q^* = \sqrt{\frac{2DS}{H}} = \sqrt{\frac{2\times 250{,}000\times 150}{12}} = 2500 \quad (4)$$

Total Cost (TC) at EOQ:

$$TC(Q^*) = S \times \frac{D}{Q^*} + H \times \frac{Q^*}{2} = 150 \times \frac{250{,}000}{2500} + 12 \times \frac{2500}{2} = 15000 + 15000 = 30{,}000 \quad (5)$$

Optimal Number of Orders per Year:

$$\text{Optimal number of orders per year} = \frac{D}{Q^*} = \frac{250{,}000}{2500} = 100 \quad (6)$$

Length of Order Cycle Time:

$$\text{Length of order cycle time} = 250 \text{ days in a year}/100 \text{ orders} = 2.5 \text{ days} \quad (7)$$

3. Methodology

Our methodology involves the identification of characteristic profiles that can help in carrying out condition-based maintenance after the production of any single item or batch item on the production line, namely infeed centerless grinding, automated production line for tool and die making, prismatic machining center, and wire electrical discharge machine that is developed in the center. More specifically, we use the data from the development of the hard turning center in the prism lab, a new development in the prismatic machining center for the industry, centerless grinder, abandoned development of the infeed centerless grinding machine for industrial purposes, CNC turn-mill center, the dedication of the CNC lathe tending center for TAL, design, and development of a 10-station turret with milling functionality, design and development of fork and drum processing on a horizontal machining center, and anomaly detection algorithm for the spindle of the turret indexing system for a CNC lathe tending center.

All the mentioned studies had data related to noise and vibration sensor data, characterization times, lag times, tool wear estimation through acoustic emission signals, spindle speed signals, setup times for feature extraction modules, and mean detection delay for the analysis, feature extraction module for roughness, tool wear, feature extraction through sub-band envelope analysis, cyclostationary signal analysis, fast Fourier transform, discrete wavelet transform-correlation function, pulse-coupled neural network, Karl Pearson's correlation coefficient, skewness, and wavelet-based feature extraction module. The studies also comprised laptop parameter tool wear sign prediction model development for in-line condition-based preventive maintenance for the spindle of the CNC lathe tending center, bearing parameter tool wear sign prediction model development for in-line condition-based preventive maintenance for the spindle of the turret indexing system for a CNC lathe tending center, training of neural networks, numerical validation of the results, developed setups, tool wear monitoring and forecasting models design, and results.

3.1 Data Collection and Preprocessing

Data collection in manufacturing refers to the extraction and gathering of data from numerous sources. Data is collected from a variety of sources such as equipment sensors and outputs, human inputs or outputs, operational data, and the factory network. It is important to collect clear data to extract patterns and knowledge. Thus, sensor installation in an enterprise, as well as process parameter signaling, is critical. Problems can arise from numerous sensors sharing the same parameter signal of a manufacturing process and a supervisor not having visibility into obscured production quality.

The data for the training of machine learning algorithms comes from various sensors located on the machines as shown in figure 3. All components of a sensor must be maintained at certain intervals to ensure clear data results used by RF and ML models. Thus, it is a concerning task. The datasets used in a machine learning model, as well as business rules, must be found extensively. The data collected emerges in different formats, different qualities, and different time frequencies. It is important to have a well-formatted dataset for the optimized model. In conclusion, a rule-based structure is required in manufacturing to start the collaboration of AI/ML applications.

Figure 3. Data Preprocessing

3.2 Machine Learning Model Selection

In the next step, it is necessary to pick the machine learning model or the big data analytics algorithm that best suits the needs and requirements of the manufacturing process at hand. It is important that the selected algorithm can solve the specific problem of decision support. It is possible to use machine learning for the classification of different machine tool behaviors within manufacturing, such as different tool-wear states and various data-based feature extraction and selection criteria. When comparing decision trees, ensemble methods, and sequential learning for machine-learning-based energy prediction through sensor monitoring during machining, different behaviors of these algorithms for different types of data are expected. When machine learning is used for industrial anomaly detection, various machine learning techniques can be applied.

Big data analytics, independent of industry type, could be tested when applied for root cause analysis with dynamic data. Using the advantages of deep learning, results with attractiveness for regular and continuous quality testing, process optimization, and process control are proven. This proposes a method to detect anomalies automatically online in regular and continuous time-series manufacturing systems. Machine learning methods like clustering, regression analysis, neural networks, decision trees, and random forests, along with a methodology designed for the selection of the best features in such systems, are applied to find the relations between the input and output signals. They were further used to forecast productivity and operations in the continuously running system of woven fabrics. Supervised and unsupervised learning methodologies were applied. In supervised methods, regression analysis, decision trees, random forests, and neural networks were used to predict the future status of the system. The findings show that neural networks outperform all other methods.

3.3 Implementation and Testing

Implementation of morphology detection is a complex process but can be achieved with a small number of steps and data only from SEM and SAED. First, the data are collected and used to train the best architecture that can detect the optimal morphologies. By capturing more shapes and using an up-to-date list, a model that detects a dozen morphologies of desired holes was built successfully. The trained model had been validated and tested to check the stability and, in terms of input data, five different grids were used, data volume had changed, and the number of morphologies had changed. The study successfully detected target shapes across wide ranges of acquisition settings, acquisition time, and data size. Providing a very simple configuration, it is important to note that the detected shape database is the layer's quality driver, practically representing the scanner's automatic daily setup.

4. Case Studies and Applications

Title: 4.1 Steel Captive Power Plants Optimization Case Summary: 46 Steel Captive Power Plant (CPP) plants ranging from 4 MW to 400 MW were used as the case for energy efficiency and operation optimization. The data comes from 46 power plants with a combined capacity of 5 GW, 200,000 production data points, plus 1 billion operational logs. Tools of the Energy Management Data Platform using big data analytics and artificial intelligence technologies were developed and applied to achieve a high standard of sustainable energy use, efficient production, and reliable power supply. Value Proposition: The solution in service by a company achieved more than 10% energy efficiency improvement in the CPP operations, increased availability, and reduced maintenance costs. It can also offer demand-side management and predictive maintenance advisory services to utility-scale power plants in the province, autonomous territory, or self-generating industrial parks. The current deliverable service annual benefit is about 6.9 million, including increased capacity and operating power benefits. Title: 4.2 Challenges for Optimizing Wafer Manufacturing Operations Case Summary: Wafer manufacturing has a special challenge in technology due to its operational complexity and variety. This case study is concentrated on managing tens of thousands of production tools and large warehouses daily to satisfy the schedule and cycle time requirements. To fulfill the target within the acceptable operational cost, the solution has to be achieved through the application of AI/ML on process tools OEE, an AI knowledge-based warehouse system, multi-criteria scheduling, and a robust production execution system.

4.1 Real-world Examples of Production Efficiency Optimization

Solution 1: One large automotive components manufacturer provided access to the data of its MES system as well as the SCADA system. Among many possible use cases, we selected those connected with: - Identifying root causes of production losses: predictive analytics and data interpretation for analysis of production loss due to machine downtime. - Improving throughput: predictive analytics and expert rules for optimization of adjusting time. - Diagnosing problems with specific items. During the adaptation process, we found that for some equipment types, data was not reliable: e.g., a high correlation between the number of parts produced and the number of errors was found - too many errors were reported even for the records where the number of produced parts equaled zero. The delusion rate was as high as 30%. It was decided to detach this data from the data history due to the high risk of misinterpretation. The data history was quite long – about 5 years for some of the elements. The executed mission was reliable data that makes later asset measurements more informative and reliable and reduces diagnostic time. Customer benefits: The manufacturer had previously been occupied with the manual interpretation of the production losses reasons. The customer recognized that the monthly Loss Analysis reports are being delivered to the HQ. Each report was published to all involved parties in less than 2 working days after the end of the month. The response activity in production and maintenance was deployed, and the impact was visible in financials. The throughput of specific machines was also visually improved. Only 4 expert rules were provided to affect the output of the adjustments' optimizer. Solution 2: Another example related to the metal and mining industry. The plant utilized a maintenance management system, connected with condition monitoring data of the equipment. We selected the following use case: - Maintenance schedule optimization: AI/ML techniques for verification of scheduled breaks before they are done. Customer benefits: The customer requested the service as a claim reduction of high downtime losses. Only 5 basic expert rules could help to predict 2/3 of the breaks that did not require changing. The shared data model and subsequent communication with maintenance engineers made the trust in the output much stronger. Finally, the maintainers could estimate, after the first visible checks, that the system is reliable, and this directed significant input from the interest toward the Equipment Health Index. They could also request such a model for similar technically connected equipment. Several other cases for the same plant are under discussion, developing an automatic anomaly detector for the number of gears damaged in the upkeep process; and predictive modeling for kiln weight to replace real TR control with an AI/ML-based one. Improving the production efficiency based on algorithmizing of the planning process is shown in figure 4.

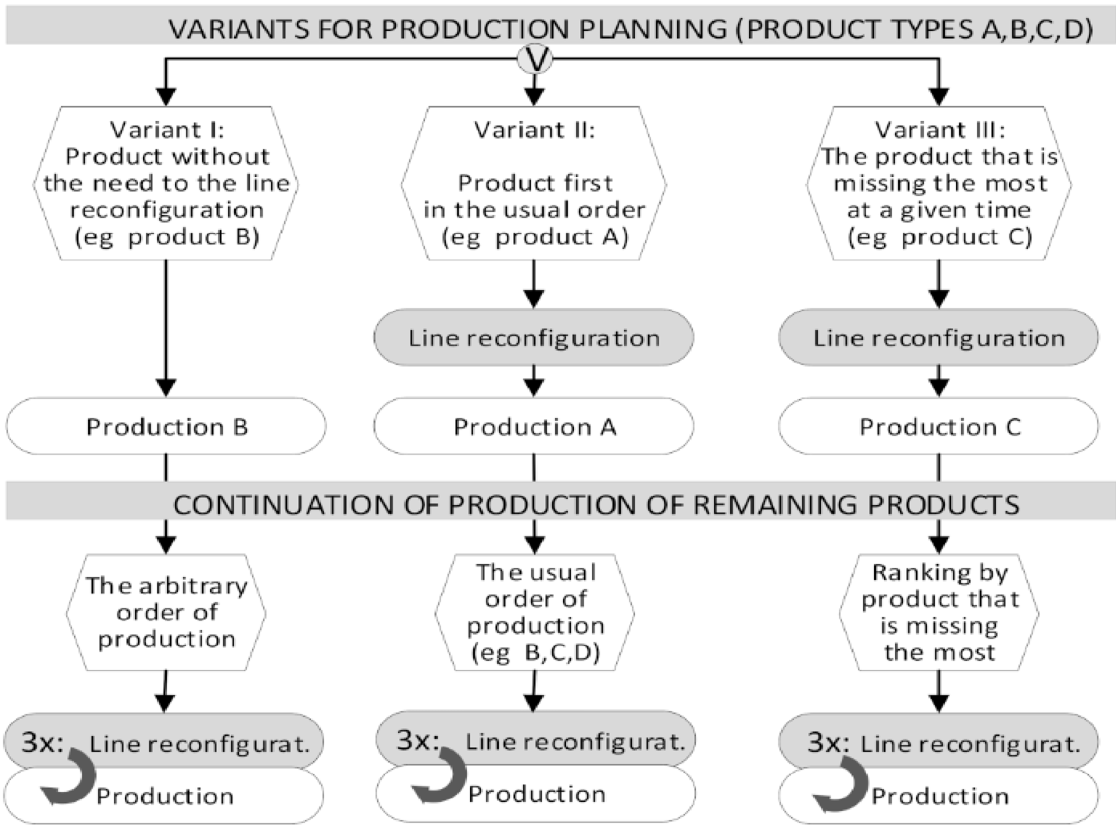

Figure 4. Improving the Production Efficiency Based on Algorithmizing of the Planning Process

4.2 Impact of Big Data and AI/ML on Key Performance Indicators

The impact of big data in combination with AI/ML on relevant KPIs will vary according to the particular sector of manufacturing. Depending on the sector, for example, the value-add of main drug product characteristics, end-product quality attributes, and related supply chain tools can be applied to achieve better production and quality performance and substantial reductions of product lead time and production cost, as are currently being experienced in semiconductor manufacturing. On the other hand, consumer manufacturing, such as food and drink, FMCG, and textiles, both bring their complex production challenges and may be handled simultaneously at multiple sites. Within the consumer goods sector of manufacturing, largely driven by retailer price pressures, the main driver of KPIs – cost, lead time, and product quality – is the factory's manufacturing performance. Optimizing production efficiency by using machine learning models can be achieved through the continuous throughput yield of high-end APCs, first-time-right of quality requirements, reduced changeover times of the production systems, enhanced use of standards and templates, and simple product shapes, colors, and recipes, without forgetting the financial constraints of typical family-owned operations. In the associated food and drink manufacturing sector, production and quality performance concerns name but a few factors important to continuous efficient operation. A common requirement is the need for acquiring data transparency from the field to the fork, requiring a combination of digital and mobile network technologies. Furthermore, to guarantee the long shelf life demanded by many products, both product and packaging integrity analyses are required to inform seal and decoration defects, thereby ensuring customer satisfaction, reducing waste, and offering brand protection.

5. Challenges and Future Directions

The last several years have seen a substantial increase in the amount and type of digital data used in manufacturing. The growth of industrial IoT in industrial settings has been accelerated by multiple technological trends and tools, including advancements in sensor technology, edge computing, ubiquitous interconnectivity, and cybersecurity. Each of these advancements is contributing to a shift within the manufacturing domain that is enabling increasingly detailed data describing manufacturing processes. While IoT and smart manufacturing are addressing issues with data accessibility for given data types, data growth has been more limited for additional contextual data types, such as factory events, work orders, production schedules, and quality and loss information. As these data streams converge in the smart manufacturing ecosystem, the breadth of the digital data may be the tipping point, redefining the manufacturing knowledge canon.

Over the last several years, the machine learning and artificial intelligence communities have invested significant effort in building models to better predict and improve different manufacturing specializations. While these models represent a significant step forward over previous approaches, there are several impediments, or "costs of predictions," preventing the realization of scalable predictions across the domain. Using big data methodologies, the manufacturing domain can be fully characterized, and high-quality feature vectors can begin to be inferred, thereby learning directly from the high-quality expert labels spanning the domain. The coupling of these communities can provide immense value across the manufacturing landscape. However, to fulfill the potential of AI/ML, particularly in a world of limited domain and engineering knowledge, manufacturing faces a new set of challenges.

5.1 Overcoming Implementation Challenges

Manufacturers should rely on a few proven strategies as they undertake their Industry 4.0 journey to accelerate manufacturing production efficiency improvements. To streamline the transformation, companies can start by narrowing the focus. Data is like water; too much can be a problem. Factories that spend years collecting every conceivable data point usually get mired in the complexity of integrating so much information. To succeed, manufacturers need a clear vision of the problems they are trying to solve and the tools they will need. Input from experienced partners can be invaluable in getting all the critical pieces aligned. Factory owners should not just rely on internal expertise or capabilities when deciding what to do with the data they can collect from previously unconnected or too-slow systems.

By homing in on the scenarios that have a clear return on investment, big data analytics can provide solutions in short order. Start small, measure results, and extend. The Industry 4.0 trend offers big potential, but not every road to automation involves a significant technological overhaul. It is better to look at it as a journey and to take a stepwise approach. The most powerful insights will often be the simplest ones. Interest in Industry 4.0 is growing, and in some cases, organizations are developing complex big data management strategies in anticipation of more powerful big data solutions. Rather than waiting for an ideal solution to be developed, however, organizations should catch the low-hanging fruit and begin benefiting from some returns as soon as possible. Since the time until payback and costs are low, it makes sense to reward the small off-the-shelf steps and execution phase work. As manufacturers embark on their Industry 4.0 journey, adopting a focused and strategic approach is crucial for accelerating production efficiency improvements. Companies should avoid the trap of collecting excessive data, which can lead to overwhelming complexity and integration challenges. Instead, they need a clear vision of the specific problems they aim to address and the tools required for solutions. Collaborating with experienced partners can help align critical elements and leverage external expertise. By targeting scenarios with a clear return on investment, manufacturers can implement big data analytics effectively, starting with small-scale initiatives that deliver measurable results. Viewing the transition as a journey allows organizations to take incremental steps rather than committing to extensive technological overhauls. The most impactful insights often stem from straightforward applications, so companies should prioritize quick wins and low-cost solutions, capitalizing on immediate benefits while laying the groundwork for more advanced big data strategies in the future.

Equation 8-12 shows the resource allocation.

Objective function is given by

Maximize:

$$a^*(n), p(n) \tag{8}$$

$\sum_{i=1}^{N}(\lambda_i(n) w_i g_{ii}(n) p_i(n) - \lambda_h h * (n))$
(8)

where

$$\sum_{j=1,j\ \ =1}^{N} g_{ij}(n)p_j(n) + \eta \tag{9}$$

Subject to constraints

1. Power constraint for each i:
   $$0 \leq p_i(n) \leq p_{i,max} \tag{10}$$
2. Interference and signal quality constraint:
   $$R_{i,min}(n) \leq \frac{w_i g_{ii}(n)p_i(n)}{\sum_{j=1,j\ \ =1}^{N} g_{ij}(n)p_j(n)+\eta} \leq R_{i,max} \tag{11}$$

3. Maximum constraint on h*(n):
   $$h^*(n) \leq h_{max} \tag{12}$$

Where

- N represents the number of entities (e.g., users, devices).
- $\lambda_i(n), w_i, g_{ii}(n), p_i(n), \lambda_h, h^*(n), p_{i,max}, R_{i,min}(n), R_{i,max}$ *and* h_{max} are parameters and variables within the optimization, with specific meanings depending on the application (e.g., power control in wireless networks, interference constraints, etc.).
- η could represent a noise term or interference margin.

5.2 Potential Advances in Technology

Potential advances in technologies include:

Machine learning pioneers in manufacturing are laying the foundation for modern solutions that look for root causes of production issues and propose systematic, scalable, and user-friendly software. Current technologies include correlation-based discovery, supervised learning, and neural network-based variable importance.

In the era of Industry 4.0, the smart data paradigm in manufacturing views data analytics as a chisel and production facilities as mosaics, both indispensable for shaping and perfecting the whole. Insights from genetic data analyses underscore a simple yet indefensible fact: data analytics, especially big data analytics, disproportionately pay off in discrete data settings because they significantly reduce the probabilities of type two prediction errors and false negatives. Manufacturers are motivated by the shared belief that business-critical problems need personal attention, but data analytics will do most of the personalization. In the context of Industry 4.0, advances in machine learning are transforming manufacturing by enabling a deeper understanding of production challenges through innovative data analytics solutions. Pioneers in this field are developing systematic, scalable software that addresses root causes of issues, utilizing techniques such as correlation-based discovery, supervised learning, and neural networks to assess variable importance. The smart data paradigm positions data analytics as a vital tool, akin to a chisel in sculpting a mosaic, highlighting the significance of big data analytics in reducing type two prediction errors and minimizing false negatives, particularly in discrete data environments. Manufacturers are increasingly recognizing the value of tailored solutions for critical business problems, with data analytics taking the lead in delivering personalized insights that drive operational efficiency and informed decision-making. Artificial intelligence in manufacturing market, by region is shown in figure 5.

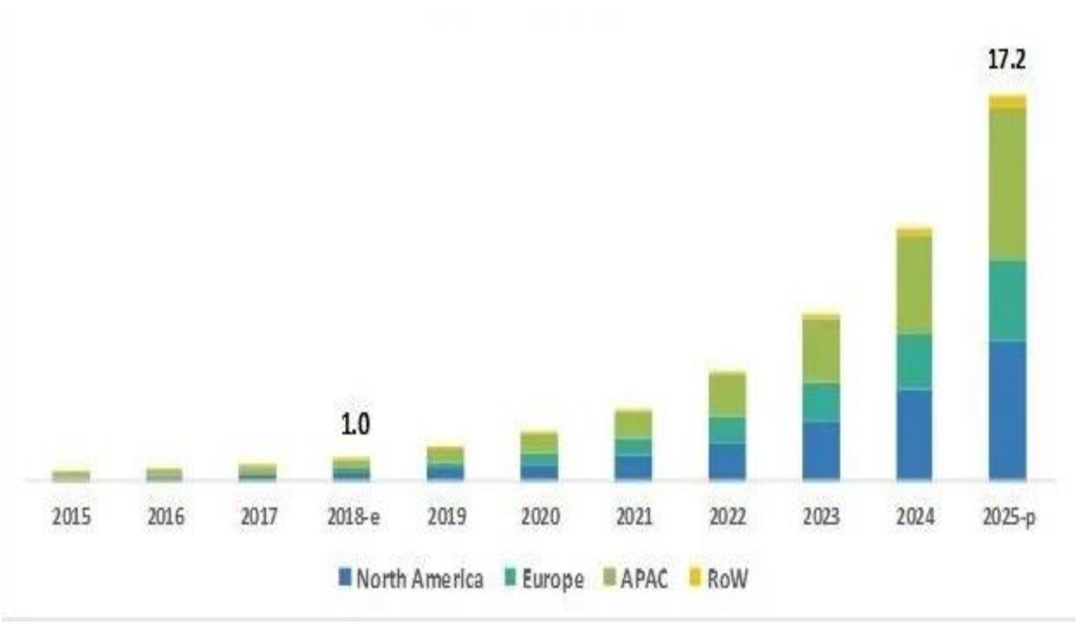

Figure 5. Artificial Intelligence in Manufacturing Market, by Region (USD billion)

6. Conclusion

The predictive and proactive value of Big Data and AI driven by sophisticated algorithms has the potential to address a multitude of issues facing computer-integrated manufacturing and the transformation to the goals of Industry 4.0. Aside from the descriptive and diagnostic applications for Big Data, it also can anticipate problems and prescribe solutions. While every manufacturing situation is unique and there are several potential complexities when embarking on the transformation to Industry 4.0, cost can be a major concern. Solutions can be crafted that work within and manage these costs, using an appropriate mix of Big Data, AI, and Machine Learning. Companies actively involved in the transformation to Industry 4.0 do indeed see a great deal of potential in the use of Big Data to handle production-related challenges. The responsive and proactive value of Big Data carried out by sophisticated algorithms has the potential to contribute to solving a very large number of challenges with which manufacturers struggle and to help users realize the real benefits of Industry 4.0. Companies are adapting their infrastructure and data methodologies to capitalize on the potential of Big Data using increasingly sophisticated software solutions that provide in-depth processing in a reasonable amount of time and at a manageable cost.

6.1 Future Trends

In a recent survey of manufacturing executives, the focus shifted to trying to capture and exploit the big data being generated on shop floors by equipment and the vast amounts of data from processes that had been received. The study also reports that predictive maintenance systems using IoT sensor data have emerged as the top use case for companies that have already implemented these types of big data services. Finding improvement opportunities faster and making those benefits more visible is compelling. Companies that are leaders in reducing downtime or improving efficiency do so through optimizations that enable them to work smarter to produce more. They often face similar trade-offs and have to consider similar constraints. They have complex configurations of equipment and are equally challenged by process models that cannot capture everything about their equipment configurations and still be quickly and accurately solved. They require similarly ingenious ways of approximating an optimal solution in a reasonable amount of time.

This study provides a framework for observation and measurement that increases confidence in a predictive algorithm's inherent knowledge representation and capability as needed in various stages of optimization. Though the example used is that of a specific type of data-producing equipment, the executives interviewed pointed out that the need for increased visibility and accuracy is also pertinent for other equipment and processes on the shop floor and supply chain partners. A similar approach could be taken to other manufacturing use cases that are important and have value before new equipment upgrades. Hopefully, a framework for seeding future development will be provided to practitioners who wish to leverage big data IoT and big data AI/ML tools to help them accelerate progress toward digital transformation ambitions – in days rather than years.

References

1. Kumar Vaka Rajesh, D. (2024). Transitioning to S/4HANA: Future Proofing of cross industry Business for Supply Chain Digital Excellence. In International Journal of Science and Research (IJSR) (Vol. 13, Issue 4, pp. 488–494). International Journal of Science and Research. https://doi.org/10.21275/sr24406024048
2. Pillai, S. E. V. S., Avacharmal, R., Reddy, R. A., Pareek, P. K., & Zanke, P. (2024, April). Transductive–Long Short-Term Memory Network for the Fake News Detection. In 2024 Third International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE) (pp. 1-4). IEEE.
3. Zanke, P., Deep, S., Pamulaparti Venkata, S., & Sontakke, D. Optimizing Worker's Compensation Outcomes Through Technology: A Review and Framework for Implementations.
4. Mahida, A. Secure Data Outsourcing Techniques for Cloud Storage.
5. Chintale, P., Deshmukh, H., & Desaboyina, G. Ensuring regulatory compliance for remote financial operations in the COVID-19 ERA.
6. Vaka, D. K. (2024). Procurement 4.0: Leveraging Technology for Transformative Processes. Journal of Scientific and Engineering Research, 11(3), 278-282.
7. Manukonda, K. R. R. Multi-User Virtual Reality Model for Gaming Applications using 6DoF.
8. Manavadaria, M. S., Mandala, V., Surabhi, S. N. R. D., Manoharan, S., Gupta, R., & Londhe, P. M. (2024, July). Smart City Traffic Monitoring and Control: Integrating Wireless Sensors with KNN-TCGAN Model. In 2024 International Conference on Data Science and Network Security (ICDSNS) (pp. 1-6). IEEE.
9. Kommisetty, P. D. N. K., & Nishanth, A. (2024). AI-Driven Enhancements in Cloud Computing: Exploring the Synergies of Machine Learning and Generative AI. In IARJSET (Vol. 9, Issue 10). Tejass Publishers. https://doi.org/10.17148/iarjset.2022.91020
10. Vaka, D. K. (2024). The SAP S/4HANA Migration Roadmap: From Planning to Execution. Journal of Scientific and Engineering Research, 11(6), 46-54.
11. Avacharmal, R. (2024). Explainable AI: Bridging the Gap between Machine Learning Models and Human Understanding. Journal of Informatics Education and Research, 4(2).
12. Pamulaparti Venkata, S., & Avacharmal, R. (2023). Leveraging Interpretable Machine Learning for Granular Risk Stratification in Hospital Readmission: Unveiling Actionable

Insights from Electronic Health Records. Hong Kong Journal of AI and Medicine, 3(1), 58-84.
13. Mahida, A., Chintale, P., & Deshmukh, H. (2024). Enhancing Fraud Detection in Real Time using DataOps on Elastic Platforms.
14. Chintale, P., Korada, L., WA, L., Mahida, A., Ranjan, P., & Desaboyina, G. Risk Management Strategies for Cloud-Native Fintech Applications During the Pandemic.
15. Muthu, J., & Vaka, D. K. (2024). Recent Trends In Supply Chain Management Using Artificial Intelligence And Machine Learning In Manufacturing. In Educational Administration Theory and Practices. Green Publication. https://doi.org/10.53555/kuey.v30i6.6499
16. Manukonda, K. R. R. (2024). Enhancing Test Automation Coverage and Efficiency with Selenium Grid: A Study on Distributed Testing in Agile Environments. Technology (IJARET), 15(3), 119-127.
17. Mandala, V., & Mandala, M. S. (2022). ANATOMY OF BIG DATA LAKE HOUSES. NeuroQuantology, 20(9), 6413.
18. Kommisetty, P. D. N. K., & Abhireddy, N. (2024). Cloud Migration Strategies: Ensuring Seamless Integration and Scalability in Dynamic Business Environments. In the International Journal of Engineering and Computer Science (Vol. 13, Issue 04, pp. 26146–26156). Valley International. https://doi.org/10.18535/ijecs/v13i04.4812
19. Vaka, D. K. (2024). Integrating Inventory Management and Distribution: A Holistic Supply Chain Strategy. In the International Journal of Managing Value and Supply Chains (Vol. 15, Issue 2, pp. 13–23). Academy and Industry Research Collaboration Center (AIRCC). https://doi.org/10.5121/ijmvsc.2024.15202
20. Avacharmal, R., Pamulaparthyvenkata, S., & Gudala, L. (2023). Unveiling the Pandora's Box: A Multifaceted Exploration of Ethical Considerations in Generative AI for Financial Services and Healthcare. Hong Kong Journal of AI and Medicine, 3(1), 84-99.
21. Pamulaparti Venkata, S. (2023). Optimizing Resource Allocation for Value-Based Care (VBC) Implementation: A Multifaceted Approach to Mitigate Staffing and Technological Impediments towards Delivering High-Quality, Cost-Effective Healthcare. Australian Journal of Machine Learning Research & Applications, 3(2), 304-330.
22. Mahida, A. (2024). Integrating Observability with DevOps Practices in Financial Services Technologies: A Study on Enhancing Software Development and Operational Resilience. International Journal of Advanced Computer Science & Applications, 15(7).
23. Chintale, P., & Desaboyina, G. (2018). Flux: Automating Cluster State Management and Updates Through Gitops in Kubernetes. International Journal of Innovation Studies, 2(2).
24. Vaka, D. K., & Azmeera, R. Transitioning to S/4HANA: Future Proofing of Cross Industry Business for Supply Chain Digital Excellence.
25. Manukonda, K. R. R. (2024). Analyzing the Impact of the AT&T and Blackrock Gigapower Joint Venture on Fiber Optic Connectivity and Market Accessibility. European Journal of Advances in Engineering and Technology, 11(5), 50-56.
26. Kommisetty, P. D. N. K., & dileep, V. (2024). Robust Cybersecurity Measures: Strategies for Safeguarding Organizational Assets and Sensitive Information. In IJARCCE (Vol. 13, Issue 8). Tejass Publishers. https://doi.org/10.17148/ijarcce.2024.13832
27. Vaka, D. K. (2024). From Complexity to Simplicity: AI's Route Optimization in Supply Chain Management. In Journal of Artificial Intelligence, Machine Learning and Data Science (Vol. 2, Issue 1, pp. 386–389). United Research Forum. https://doi.org/10.51219/jaimld/dilip-kumar-vaka/100

28. Avacharmal, R., Sadhu, A. K. R., & Bojja, S. G. R. (2023). Forging Interdisciplinary Pathways: A Comprehensive Exploration of Cross-Disciplinary Approaches to Bolstering Artificial Intelligence Robustness and Reliability. Journal of AI-Assisted Scientific Discovery, 3(2), 364-370.
29. Pamulaparti Venkata, S., Reddy, S. G., & Singh, S. (2023). Leveraging Technological Advancements to Optimize Healthcare Delivery: A Comprehensive Analysis of Value-Based Care, Patient-Centered Engagement, and Personalized Medicine Strategies. Journal of AI-Assisted Scientific Discovery, 3(2), 371-378.
30. Mahida, A. Explainable Generative Models in FinCrime. J Artif Intell Mach Learn & Data Sci 2023, 1(2), 205-208.
31. Chintale, P., Khanna, A., Korada, L., Desaboyina, G., & Nerella, H. AI-Enhanced Cybersecurity Measures for Protecting Financial Assets.
32. Vaka, D. K. Supply Chain Renaissance: Procurement 4.0 and the Technology Transformation. JEC Publication.
33. Manukonda, K. R. R. (2024). Leveraging Robotic Process Automation (RPA) for End-To-End Testing in Agile and Devops Environments: A Comparative Study. Journal of Artificial Intelligence & Cloud Computing. SRC/JAICC-334. DOI: doi. org/10.47363/JAICC/2024 (3), 315, 2-5.
34. Kommisetty, P. D. N. K., vijay, A., & bhasker rao, M. (2024). From Big Data to Actionable Insights: The Role of AI in Data Interpretation. In IARJSET (Vol. 11, Issue 8). Tejass Publishers. https://doi.org/10.17148/iarjset.2024.11831
35. Yadav, P., Prasad, S., & Vansia, R. Mortality Prediction in the ICU Utilizing Topic Model and Burstiness with Machine-Learning Techniques.
36. Avacharmal, R., Gudala, L., & Venkataramanan, S. (2023). Navigating the Labyrinth: A Comprehensive Review of Emerging Artificial Intelligence Technologies, Ethical Considerations, And Global Governance Models in the Pursuit of Trustworthy AI. Australian Journal of Machine Learning Research & Applications, 3(2), 331-347.
37. Tilala, M., Pamulaparti Venkata, S., Chawda, A. D., & Benke, A. P. Explore the Technologies and Architectures Enabling Real-Time Data Processing within Healthcare Data Lakes, and How They Facilitate Immediate Clinical Decision-Making and Patient Care Interventions. European Chemical Bulletin, 11, 4537-4542.
38. Mahida, A. (2023). Enhancing Observability in Distributed Systems-A Comprehensive Review. Journal of Mathematical & Computer Applications. SRC/JMCA-166. DOI: doi. org/10.47363/JMCA/2023 (2), 135, 2-4.
39. Perumal, A. P., & Chintale, P. Improving operational efficiency and productivity through the fusion of DevOps and SRE practices in multi-cloud operations.
40. Vaka, D. K. SAP S/4HANA: Revolutionizing Supply Chains with Best Implementation Practices. JEC Publication.
41. Raghunathan, S., Manukonda, K. R. R., Das, R. S., & Emmanni, P. S. (2024). Innovations in Tech Collaboration and Integration.
42. Kommisetty, P. D. N. K., & Nishanth, A. (2024). AI-Driven Enhancements in Cloud Computing: Exploring the Synergies of Machine Learning and Generative AI. In IARJSET (Vol. 9, Issue 10). Tejass Publishers. https://doi.org/10.17148/iarjset.2022.91020