World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(Research Article)

Check for updates

A DATA-DRIVEN FRAMEWORK FOR LEAN PRODUCTION OPTIMIZATION IN SAP-INTEGRATED MANUFACTURING SYSTEMS

Mahendrakumar Nathubhai Kalal [1,*] and O. C. Tanner [2]

[1] Department of IT, O.C.Tanner - Senior SAP Analyst.
[2] Houston, Texas, USA.
* Corresponding Author; Email: mahendrakalal14051989@gmail.com

ORCID Details
Mahendrakumar Kalal: https://orcid.org/0009-0006-4150-8937

World Journal of Advanced Research and Reviews, 2026, 31(03), 248–259

Publication history: Received on 24 July 2026; revised on 01 September 2026; accepted on 03 September 2026

Article DOI: https://doi.org/10.30574/wjarr.2026.31.3.2290

ABSTRACT

The increasing adoption of intelligent manufacturing and Industry 4.0 technologies has intensified the need for intelligent production planning and real-time operational optimization within enterprise systems. SAP Production Planning (SAP PP) and SAP Digital Manufacturing enable effective integration between production planning and shop-floor execution. However, optimizing lean production requires advanced data analytics and a reliable model for classifying operational efficiency. In this study, feature-engineering techniques were applied to an intelligent manufacturing dataset, producing 32 relevant attributes. The data were processed through outlier removal, categorical encoding, feature standardization using StandardScaler, and stratified cross-validation to ensure a robust and balanced evaluation. Comparative results showed that the Decision Tree, Naive Bayes, and Multilayer Perceptron models achieved accuracies of 94%, 91.14%, and 92%, respectively. The proposed Gradient Boosting model substantially outperformed these baseline algorithms, achieving 99.99% accuracy, precision, recall, and F1-score. These findings demonstrate the potential of ensemble learning to identify complex manufacturing patterns and support scalable, accurate, and real-time lean production optimization in SAP-enabled smart manufacturing environments.

Keywords: Digital manufacturing, (Systems, Applications, and Products) SAP, Production planning (PP), Lean Production Optimization, Machine Learning.

1 INTRODUCTION

The quick development of the Industry 4.0 and intelligent manufacturing has essentially changed the traditional production systems into highly interconnected and data-driven environments (Papacharalampopoulos et al. 2024; Patel 2024). Cloud-based systems, Internet of Things (IoT) and Cyber-physical systems has allowed the organizations to abandon reactive and adopt predictive and adaptive manufacturing approaches (Papacharalampopoulos et al. 2024; Patel 2024). SAP Production Planning is one of the main component of enterprise resource planning systems that allow efficient material requirement planning, capacity management, and production scheduling (Papacharalampopoulos et al. 2024). SAP PP provides alignment of the demand forecasts, inventory management and manufacturing execution processes (Papacharalampopoulos et al. 2024; Patel 2024). SAP PP offers a solid basis of operational control and efficiency by encouraging stratified planning and integration of functions with others, including procurement, sales SAP Digital Manufacturing improves the connectivity of the shop-floor by means of combining data on the machine, production orders, quality monitoring, and performance analytics into one digital platform (Papacharalampopoulos et al. 2024). It gives instant information on the performance of operations and allows organizations to make changes proactively and responsively to the changing environment of production. SAP PP is designed to support the systematic planning of activities and SAP Digital Manufacturing track the execution processes (Papacharalampopoulos et al. 2024; Patel 2024). Such integration facilitates the process of making decisions based on data and eliminates delays, deviation, and inefficiency in manufacturing systems.

Lean Production Optimization focusses on waste reduction, overproduction, reduction of downtime, and resource optimization. In lean principles, there is continual enhancement, alignment of value streams, and standardization of processes in order to increase their efficiency in operation (Papacharalampopoulos et al. 2024; Patel 2024). Lean plans are more quantifiable and manageable with the help of SAP-powered data visibility and analytics, allowing institu-tions to identify inefficiencies (Papacharalampopoulos et al. 2024; Patel 2024). However, more complicated analytic methods than the conventional rule-based approaches are needed to deal with the increasing complexity of production data (Papacharalampopoulos et al. 2024; Patel 2024). Machine Learning plays a critical role in Smart Manufacturing within industrial revolution 4.0 framework machine predictive models and high-technology analytics are used to maximize the performance of production. Using manufactured features that have been engineered and algorithms that are based on ensemble learning, organizations can anticipate the level of efficiency, identify operational risks, and improve the accuracy of decision-making in real-time. The structure of the paper is as follows: Section II provides the background study on SAP enable smart manufacturing system, and Section III provides the research approach for this study. Section IV gives the experiment analysis and result of the model's performance. Section V conclusion and future work of study are provided.

1.1 Motivation and Contribution of the Study

The proliferation of smart manufacturing and Industry 4.0, SAP-enabled environments are increasingly in need of smart production planning and real-time optimization. In spite of the fact that SAP Production Planning and SAP Digital Manufacturing offer high integration capabilities, there are such challenges as inefficiencies, operational risks, and forecasting deviations. Conventional rule-based frameworks often fail to capture complex production patterns, whereas deep learning models can be computationally expensive and an efficient and deployable forecasting framework on production efficiency optimization. Here is a brief overview of the main contribution of this work:

- The study utilizes an intelligent manufacturing dataset integrated with SAP PP and SAP Digital Manufacturing parameters for efficiency classification.
- Implementation of a data-driven preprocessing framework including outlier removal, encoding, feature engineering, scaling, and stratified cross-validation.
- Development of an advanced Gradient Boosting model capable of capturing complex interactions among production, quality, and risk-related features for improved efficiency prediction.
- Comprehensive performance evaluation using metrics with accuracy, precision, recall, and F1-score to validate the effectiveness of the proposed optimization model.
- A scalable and deployable analytical approach that supports real-time decision-making and production optimization within SAP-enabled smart manufacturing environments.

2 LITERATURE REVIEW

This research presents a comprehensive study of intelligent manufacturing opti-mization within SAP-enabled environments, focusing on advanced ML–based classification techniques in Table I show challenges and future work. Selected relevant studies are summarized below:

Gadve and Deshmukh (2026) SAP Production Planning (PP) is an essential tool for digital transformation in manufacturing, as it helps to synchronize demand, resources, capacity, and the execution process. It also contributes to integrated business operations and sustainable growth. Improved operational efficiency, reduced lead times, optimized resource utilisation, and sustainability outcomes through arete reduction and energy-efficient planning are some of the benefits of integrating SAP PP with other modules, such as SAP MM, SD, QM, and FI/CO. Other emerging technologies to benefit from this integration include digital twins and predictive analytics (Papacharalampopoulos et al. 2024). Another study by K. Immidi (2025) transformed supply chain management with the implementation of an AI-powered SAP Integrated Business Plan. AI-enhanced IBP changes the way traditional problems like demand forecasting, inventory optimization, forecast error risk, minimum inventory carrying cost, and exponential risk mitigation in supply chains are handled. Machine learning algorithms primarily help with demand forecasting, inventory management, and risk mitigation, which benefits firms with resilient and optimized supply chains (Papacharalampopoulos et al. 2024). Vaidya (2025) a new outline that integrates Digital Twin (DT) technology with SAP S/4HANA of raising real-time visibility, forecasting accuracy, and responsiveness in manufacturing environment for enterprise resource planning and production planning, integration of DT with SAP enables proactive rescheduling, predictive maintenance acts, and adaptive material requirements planning (MRP) robustness enhanced uncover transformative improvements in forecasting accuracy (up to 18%) to legacy planning approaches for organizations that desire intelligent self-correcting supply chain networks (Papacharalampopoulos et al. 2024). Q. Fu, (2025) are leading to overproduction rates of 15-22%, delivery delays of 20-30%, and cost wastage of 12-18% as a result of SAP systems and lean production (LP). Integrating generative AI into SAP modules, together with a Real-Time Lean Rule Embedding Engine (RTLREE) and TPM rules, has resulted in a rule execution accuracy of 99.2%, and is anticipated to significantly decrease maintenance cost (Papacharalampopoulos et al. 2024).

Chenna (2024) explore SAP Production Planning (PP), describing how cutting-edge technologies like ML, AI, and robotic process automation (RPA) are integrated, and Process Mining is revolutionizing manufacturing operations hyperautomation, its significant effects on efficiency, accuracy, and real-time insights in production planning, and the specific applications of each technology within the SAP PP ecosystem substantial improvements in productivity, cost reduction, and decision-making accuracy (Papacharalampopoulos et al. 2024). Lastly Bachchan (2023) AI-driven solutions, such as SAP Digital Manufacturing Cloud (DMC) and SAP Business AI, to enhance operational efficiency and product quality. These innovations leverage real-time data analytics, predictive maintenance, and intelligent automation to optimize production AI-powered quality management systems enable early detection of defects, reducing inspec-tion. Manufacturing agility, compliance, and customer happiness are all greatly enhanced by ML algorithms that analyze previous data to predict equipment failures. This enables proactive maintenance and minimizes downtime, ultimately leading to remarkable benefits.

2.1 Research gap

Table 1: Comparative analysis of SAP-Enabled Manufacturing Optimization using machine learning techniques

Author	Dataset	Key Findings	Methodology	Challenges	Future Work
Gadve & Deshmukh (2026)	SAP PP in manufacturing digital transformation	SAP PP enhances integration with MM, SD, QM, FI/CO; improves operational efficiency, reduces lead time, optimizes resources, and supports sustainability	Conceptual and system integration analysis of SAP PP with business modules	Cross-functional system complexity and digital transition barriers	Deeper integration with Digital Twins and predictive analytics
Kesava Immidi (2025)	SAP IBP with AI for supply chain management	ML improves demand forecasting, reduces inventory cost, enhances risk mitigation and supply chain resilience	AI-driven forecasting and optimization framework within SAP IBP	Forecast uncertainty and supply chain risk management	Advanced AI-driven autonomous supply chain planning
Vaidya (2025)	Digital Twin integrated with SAP S/4HANA	DT integration improves real-time visibility and forecasting accuracy (up to 18%)	Digital Twin framework combined with SAP ERP & MRP systems	Implementation complexity and legacy system adaptation	Intelligent self-correcting supply chain networks
Q. Fu (2025)	SAP + Lean Production integration	Lean-SAP misalignment causes 15–22% overproduction; RTLREE achieved 99.2% rule execution accuracy	Real-Time Lean Rule Embedding Engine (RTLREE) integrated with SAP modules	Lean rule integration complexity; operational inefficiencies	Generative AI integration to reduce maintenance cost
Chenna (2024)	SAP PP with RPA, AI, ML & Process Mining	Hyper automation improves efficiency, accuracy, and real-time insights in production planning	Integration of AI, ML, RPA, and process mining within SAP PP ecosystem	Implementation cost and system integration challenges	Expansion of AI-based intelligent decision support
Bachchan (2023)	SAP Digital Manufacturing Cloud (DMC) & SAP Business AI	AI-driven predictive maintenance reduces downtime; improves quality and agility	Real-time analytics, ML-based predictive maintenance & automation	Data integration and system scalability	Enhanced AI-powered quality management systems

Whereas the existing studies show that SAP integration, AI forecasting, and Digital Twin applications are important, few studies have proposed a combined framework of SAP PP, SAP Digital Manufacturing, and advanced ensemble learning to achieve real-time efficiency optimization. There is a lack of an integrated data driven method for accurately predicting production performance.

3 METHODOLOGY

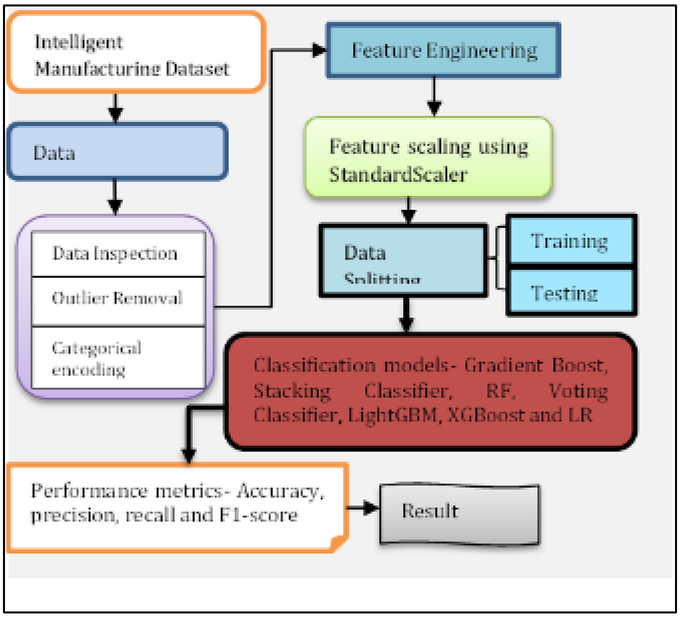


Figure 1: Flowchart for SAP production planning using machine learning.

The process begins with collecting data from the intelligent manufacturing environment integrated with SAP systems, including Digital Manufacturing shown in Fig. 1. The data are subjected to preprocessing namely, data inspection, box plot analysis to remove outliers, and categorical encoding. Engineering of features is then conducted to come up with meaningful operational and risk-based variables. Numerical features are normalized using StandardScaler. The processed data are then split into training and testing sets through stratified sampling to preserve the Low, Medium, and High efficiency classes. Gradient Boosting, Stacking, Random Forest, Voting, LightGBM, XGBoost, and Logistic Regression classifiers are trained on the SAP-integrated manufacturing data. Their performance is evaluated in an SAP-enabled smart manufacturing context using accuracy, precision, recall, and F1-score to assess their effectiveness in optimizing production performance and supporting decision-making.

3.1 Data Collection

The Intelligent Manufacturing Dataset[1] for Predictive Optimization is a publicly available dataset from Kaggle, designed for research in smart manufacturing. AI-driven process optimization, and predictive maintenance. The dataset simulates real-time sensor data generated by industrial machine, incorporating 6G network slicing for enhanced communication and resource allocation.

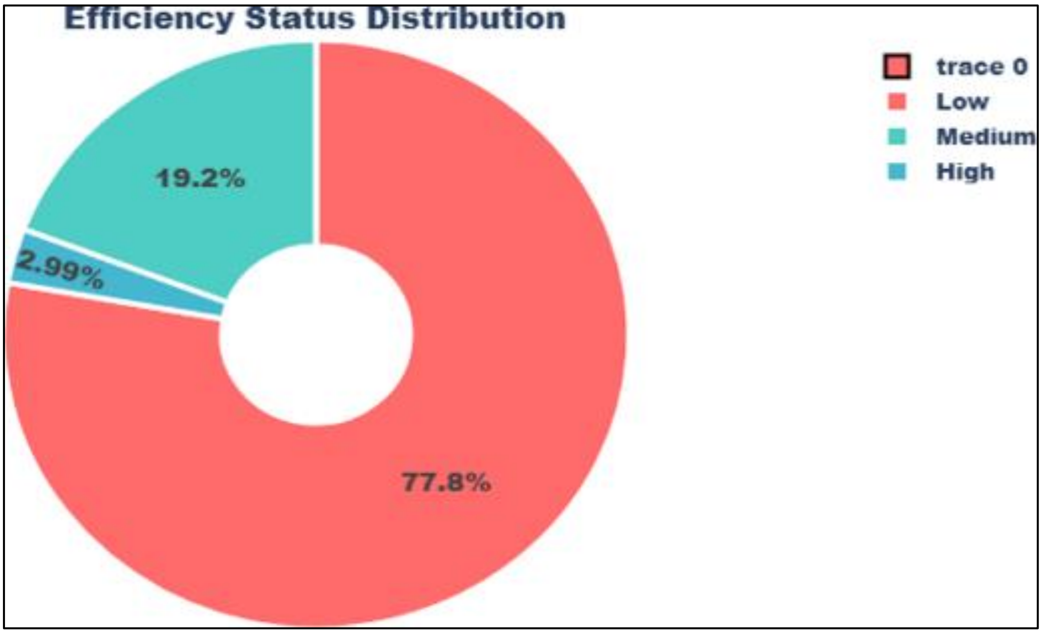


Figure 2: Donut chart distribution of efficiency status.

This donut chart shows the general distribution of the efficiency levels among the manufacturing data in Fig. 2. The majority of total instances of low (77.8) and Medium efficiency (19.2) but the High efficiency is only 2.99, to optimize production efficiency and idealize processes within the system.

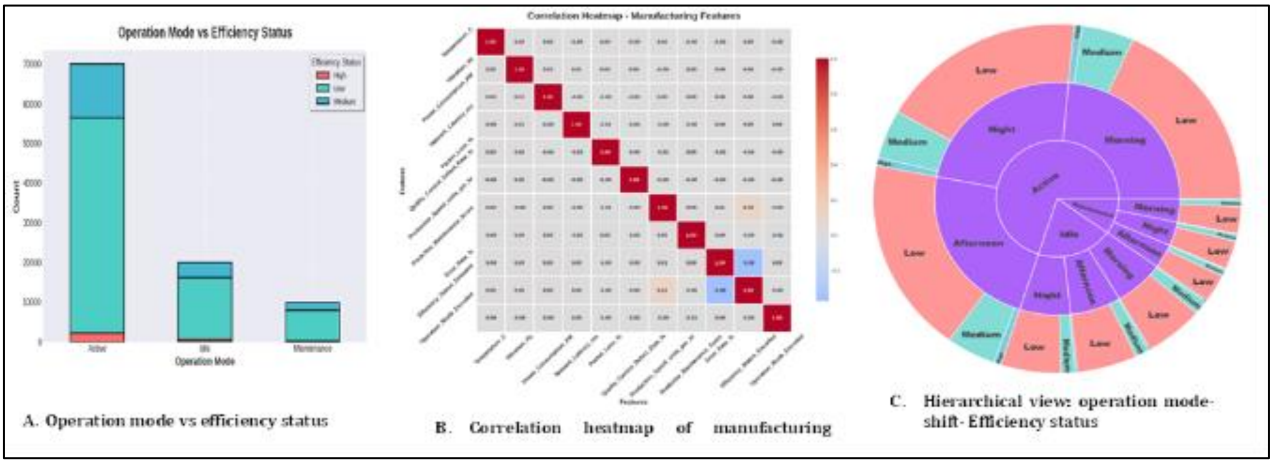


Figure 3: Manufacturing Operations Analysis: Operation Mode, Feature Correlations, and Shift-Based Efficiency Distribu

Manufacturing efficiency analysis is presented in Fig. 3. (A) A stacked bar chart Operation Mode vs. Efficiency Status shows that most operations occur in Active mode, with the majority exhibiting Medium efficiency, while Idle and Maintenance modes have significantly lower counts. (B) A correlation heatmap shows the relationships among manufacturing features, where most variables exhibit weak correlations. (C) A hierarchical sunburst chart illustrates the distribution of operation mode, shift (morning, afternoon, and night), and efficiency status, indicating that Active operations dominate across all shifts and are primarily associated with Low to Medium efficiency levels.

3.2 Data Preprocessing

Data preprocessing involved inspecting the dataset, handling missing values, re-moving outliers using box plot, and encoding categorical variables into numerical format. Feature engineering and standard scaling then applied before splitting the data using stratified cross-validation to ensure reliable and balanced model training are described below:

- **Data inspection:** Data inspection is performed to view the structure of data, type of features and statistics distribution. Preprocessing and model development are doing after checking the data to verify quality and reliability, including missing values, duplication, and inconsistency.

- **Outlier removal:** Box plot analysis is used to remove outliers of numerical features. The data that are beyond the range of the whisker (determined by the Interquartile Range, IQR) were viewed as outliers. This approach assisted in identifying abnormal production measures, the enhancement of the consistency of data.
- **Categorical encoding:** Categorical encoding is used to convert non-numerical variables into numerical format that can be used in machine learning algorithms. During the implementation, categorical characteristics like the mode of operations are converted with the help of the encoding methods, whereas the target variable (Low, Medium, High) are converted into the numerical labels.

3.3 Feature Engineering

Feature engineering is carried out to make the model have greater predictive ability, by creating meaningful attributes out of the available manufacturing data. It is performed by creating new derived attributes such as Temperature_Vibration_Ratio, Network_Quality_Score, High_Error_Risk, High_Defect_Risk, Temp_Power_Interaction, and Error_Defect_Combined to capture operational and risk-related patterns. Along with core features like Power_Efficiency, Qual-ity Score, and Reliability_Score, so the total features increased to 32, enhancing model performance as shown in Table II.

Table 2: Advanced Feature Engineering

Index	Power_Efficiency	Overall_Quality_Score	Reliability_Score	High_Temperature_Risk	Efficiency_Status
0	49.69	77.28	0.29	0	Low
1	121.82	87.33	0.71	1	Low
2	15.13	91.34	0.91	0	Low
3	65.03	92.68	0.96	0	Medium
4	22.02	85.61	0.50	1	Low
5	13.59	95.24	0.90	1	Low
6	61.65	86.07	0.20	1	Low
7	27.88	95.03	0.91	0	Medium
8	40.35	81.00	0.44	1	Low
9	63.90	89.98	0.78	0	Medium

3.4 Feature scaling using StandardScaler

The StandardScaler method are used to perform feature scaling and standardized the numerical features by eliminating the mean and bringing them to unit variance. This makes sure that every feature has equal contribution in the model training and there not be the bias of those with high values. The transformation follows in Equation (1):

$$Z = \frac{E-\mu}{\sigma} \quad (1)$$

Where denotes the standard deviation and represents the mean

3.5 Data Splitting

The dataset is split into 20% testing and 80% training sets using a fixed random state. Stratified K-Fold cross-validation (k = 5) are applied to ensure balanced representation of Low, Medium, and High efficiency categories in each fold.

3.6 Classification Gradient boosting for digital manufacturing

A supervised machine learning algorithm is the Boosting Technique. The boosting approach is an ensemble strategy that uses a large number of weak learners that concentrate on the mistakes that happen at each stage until a strong model for regression and classification is generated using a parameter of random_state=42 and no estimators=100. The difference between actual and expected values is found using the loss function. A decision tree for prediction is a weak learner. A model that minimizes the loss function by adding. To improve the model's forecast and reduce its prediction error, each weak learner tries to correct mistakes made by earlier weak learners. Examine a collection of arbitrary input variables.

={ 1, 2,… } in addition to a response variable in Equation (2):

$\boldsymbol{\Gamma}(x) = a\tau g \min_{\Gamma(x)} L_{x,x}(x, \boldsymbol{\Gamma}(x))$ (2)

The approximation function in Equation (3) is estimated using the loss function as a squared error function.

$Loss(x, \boldsymbol{\Gamma}(x)) = (x - \boldsymbol{\Gamma}(x))^2$ (3)

The gradient of the loss function (, ( )) in Equation (4) may be obtained using the following: $_i$

$$\hat{z_i} = \left[\frac{\partial Loss(z_i, F(x_i))}{\partial F(x_i)}\right]_{F(x)=F_{m-1(x)}}$$

It is possible to generalize the gradient's computation range while employing regression trees $h(xi;b)$ with parameter b as weak learners. In most cases, it takes x as an input variable and returns b as a parameterized function. To obtain the tree, one may solve the following Equation (5).

$$b_m = arg \min_{b,\beta} [\hat{z}_i - \beta h(x_i, b)]^2$$

The parameters acquired at iteration u and the weight value, bm, which is also called the expansion coefficient for every student who is weak, are represented by m and β, respectively.

3.7 Performance Metrics

The performance of production planning and digital manufacturing system used various metrics, in addition to the loss, the following four values may be obtained: true positive (TP), false positive (FP), false negative (FN), and true negative (TN). These values include accuracy (AC), precision (PR), recall (R), and F1-score (F1) via confusion matrix. The performance of every classifier is measured through the following metrics in Equations (6) to (9).

$$Accuracy = \frac{\text{TN + TP}}{\text{TP + TN + FP + FN}} \quad (6)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (7)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad (8)$$

$$\text{F1} = \frac{2*(\text{precision}*\text{recall})}{\text{precision+recall}} \quad (9)$$

The GB model confusion matrix indicates that the model has a good classification rate shown in Fig. 4. It rightly forecasts 597 High, 15,565 Low and 3836 Medium. The numbers of misclassification of low cases as medium are only 2. The classification performance of proposed model. A prec, rec and F1of 1.00 on all classifications: High (597 samples), Low (15,565 samples) and Medium (3,838 samples). The total accuracy achieved is 1.00 in 20,000 total instances. The weighted average score is 1.00 which means that perfect classification.

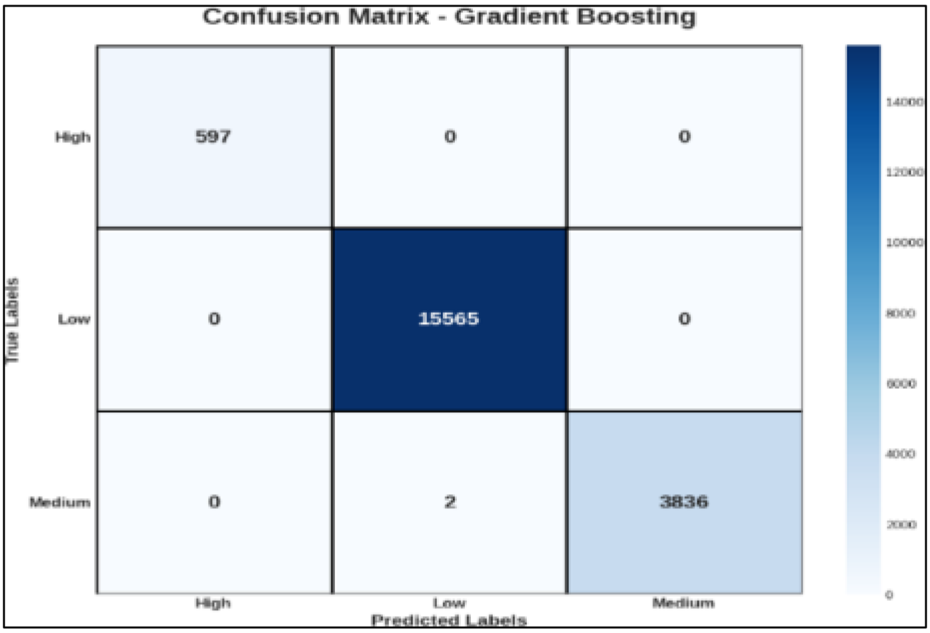
The metrics collectively provide insights into the model's accuracy and efficiency in predicting the target variable.

4 RESULT AND DISCUSSION

The proposed model for lean production optimization is experimentally evaluated on the Intelligent Manufacturing dataset using Google Colab with 4GB RAM and 256GB SSD in Table III. The model's performance is evaluated using the F1, acc, prec, and rec. The proposed Gradient Boosting model outperformed all of the other models in terms of accuracy, prec, rec, and F1, reaching 99.99%. The results show that the suggested method for combining SAP PP and SAP Digital Manufacturing is effective in improving intelligent production optimization.

Table 3: Performance model on intelligent manufacturing dataset

Evaluation Metrics	Accuracy	Precision	Recall	F1-score
Gradient boosting	99.99	99.99	99.99	99.99
Stacking classifier	99.98	99.98	99.98	99.98
Random forest	99.98	99.98	99.98	99.98
Voting classifier (Hybrid)	99.95	99.95	99.95	99.95
LightGBM	99.86	99.86	99.86	99.85
XGBoost	99.79	99.79	99.79	99.79
Logistic Regression	91.64	91.49	91.64	91.55




```
==========================================================
CLASSIFICATION REPORT - Gradient Boosting
==========================================================
              precision    recall  f1-score   support

        High       1.00      1.00      1.00       597
         Low       1.00      1.00      1.00     15565
      Medium       1.00      1.00      1.00      3838

    accuracy                           1.00     20000
   macro avg       1.00      1.00      1.00     20000
weighted avg       1.00      1.00      1.00     20000
```

Figure 4: Confusion matrix and classification report of Gradient boosting.

The GB model confusion matrix indicates that the model has a good classification rate shown in Fig. 4. It rightly forecasts 597 High, 15,565 Low and 3836 Medium. The numbers of misclassification of low cases as medium are only 2. The classification performance of proposed model. A prec, rec and F1of 1.00 on all classifications: High (597 samples), Low (15,565 samples) and Medium (3,838 samples). The total accuracy achieved is 1.00 in 20,000 total instances. The weighted average score is 1.00 which means that perfect classification.

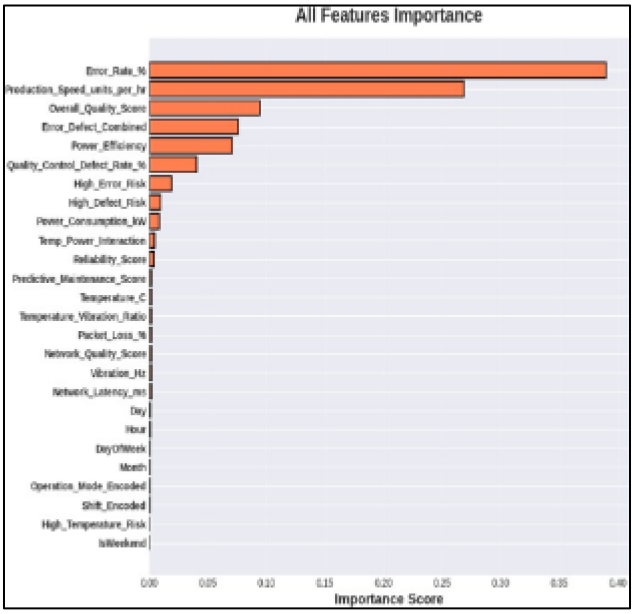


Figure 5: Feature Importance Aligned with Lean Principles.

Fig. 5 shows the relative importance scores of the 32-feature used in the model. Error_Rate_% has the highest importance (approximately 0.39), followed by Production_Speed_units_per_hr (~0.28). Features such as Overall_Qual-ity_Score, Error_Defect_Combined, and *Power_Efficiency* show moderate influence, while variables like IsWeekend, Shift_Encoded, and High_Tempera-ture_Risk contribute minimally.

4.1 Fold Cross Validation Result

The cross-validation performance of the evaluated machine learning models in Table IV. The stacking classifier achieved the highest mean cross-validation score of 0.999938 with minimal standard deviation, indicating highly stable and accurate predictions. Random Forest also demonstrated strong performance with a mean score of 0.999900. The remaining models, including Voting Classifier, LightGBM, and XGBoost, achieved slightly lower.

Table 4: Cross validation summary result

Model	Mean CV Score	Std CV Score	Min CV Score	Max CV Score
Stacking Classifier	0.999938	0.000125	0.999687	1.000000
Random Forest	0.999900	0.000075	0.999812	1.000000
Voting Classifier	0.999638	0.000320	0.999188	0.999938
LightGBM	0.998500	0.000535	0.997875	0.999437
XGBoost	0.998475	0.000639	0.997500	0.999375

4.2 Comparative Analysis

The comparison between the performance of various models to lean production optimization in an integrated system of SAP and Digital Manufacturing shown in Table V. The conventional machine learning methods like DT (94.23%), NB (91.14%), and MLP (92%) performed well but relatively poorer in acc, prec, rec, and F1. Conversely, the proposed Gradient Boosting model is much more successful than all the baseline techniques with an acc, prec, rec, and F1 of 99.99. These findings indicate that the superior ensemble learning methods can be effective in improving the quality of production planning, increasing operational stability, and making data-driven decisions in intelligent manufacturing systems.

The resulting strategy employs a Gradient Boosting algorithm to enhance predic-tive accuracy and operational efficiency within an integrated production planning and digital manufacturing framework. This ensemble technique is useful to in-crease the accuracy, recall, and performance of the overall classification. This model is useful in analyzing the characteristics of production, minimizing the de-viation of planning, and data-driven decision-making, which enhances reliability in manufacturing decision-making and optimizes the results of the processes in smart industrial systems.

Table IV: Comparative Analysis of Models for Lean Production Optimization in SAP

Measure	Accuracy	Precision	Recall	F1-score
DT[23]	94.23	95.1	94.2	94.2
NB[24]	91.14	91.8	91.1	91.1
MLP[25]	92	85	92	88
Gradient boosting	99.99	99.99	99.99	99.99

4.3 Justification and Novelty

This research is justified by the growing need to ensure smart and adaptive production optimization in SAP-powered smart manufacturing systems. In spite of the progress in SAP Production Planning and SAP Digital Manufacturing, the current solutions to the problem are mostly based on the use of rule-driven logic or distinct AI solutions that fail to reflect all the intricate interdependences between operational, quality, and risk-related parameters. The novelty of data-driven model that integrates features engineering and an ensemble Gradient Boosting model coupled with stratified cross-validation. This integrated solution improves predictability, scalability, and real-time decision-making to optimize manufacturing efficiency.

5 CONCLUSION AND FUTURE WORK

This study proposes a unified intelligent system for lean production optimization based on the integration of SAP Production Planning and SAP Digital Manufacturing systems with the most modern techniques of machine learning. The approach proposed included extensive data preprocessing, feature engineering, standard scaling, and stratified cross-validation to guarantee the quality, strength, and limited assessment of the data. The Gradient Boosting model accuracy 99.99% in identifying complicated interdependencies in production processes by using an amended manufacturing data set operational, quality, and risk-related numbers. Comparative analysis ensemble model performed better than the traditional methods its appropriateness in the correct classification of the efficiency and the support of decisions based on data. SAP-based real-time manufacturing execution, combined with the planning information, helps to improve the level of transparency of the operations, reduce the inefficiencies, and integrate the performance of production with lean manufacturing. The framework can be scaled to future work with the inclusion of real-time IoT sensors feeds, Digital Twin models, and adaptive optimization algorithms like reinforcement learning. Furthermore, large-scale industrial validation on various manufacturing fields and implementation on cloud-based environments of SAP would even extend the scalability, interoperability, and real-world applicability.

STATEMENTS ON COMPLIANCE WITH ETHICAL STANDARDS

Acknowledgments

The author acknowledges the publicly available software tools and computational resources used to support the data analysis and preparation of this manuscript.

Disclosure of conflict of interest

The author declares no conflict of interest regarding this manuscript.

REFERENCES

[1] M. A. Rahman, J. Bhowmik, M. S. Ahamed, and R. Rahman, “Opportunities and Challenges in Data Analysis Using SAP: A Review of ERP Software Performance,” *Glob. Mainstream J.*, vol. 1, no. 4, pp. 50–67, Aug. 2024, doi: 10.62304/ijmisds.v1i04.192.

[2] S. Selvaraj, “AI-Powered SAP for Smart Manufacturing: Enhancing Efficiency with Intelligent Automation,” *Int. J. AI, BigData, Comput. Manag. Stud.*, vol. 6, no. 1, pp. 78–88, 2025, doi: 10.63282/3050-9416.ijaibdcms-v6i1p108.

[3] R. Snehamrutha, “Data-Driven Optimization of Pharmaceutical Manufacturing Processes using Quality by Design (QbD) Frameworks,” *Int. J. Curr. Eng. Technol.*, vol. 14, no. 6, pp. 557–566, 2024.

[4] S. Singamsetty and S. Sanakkayala, “Agentic AI-driven portfolio optimization: a hybrid approach for optimized stock selection and deep learning in algorithmic trading,” *Multimed. Syst.*, vol. 32, no. 1, p. 70, 2026, doi: 10.1007/s00530-025-02131-7.

[5] T. L. James, D. F. Cook, S. Conlon, K. B. Keeling, S. Collignon, and T. White, “A framework to explore innovation at SAP through bibliometric analysis of patent applications,” *Expert Syst. Appl.*, vol. 42, no. 24, pp. 9389–9401, Dec. 2015, doi: 10.1016/j.eswa.2015.08.007.

[6] J. W. Sajja, “AI-Powered Hyperautomation in SAP S/4HANA Migration: Transforming ERP Transitions,” vol. 13, no. 3, p. 28, 2025.

[7] A. Nerella, N. Kolli, and J. W. Sajja, “Building Secure AI Agents for Autonomous Data Access in Compliance/Regulatory-Critical Environments,” *SSRN Electron. J.*, p. 11, 2025, doi: 10.2139/ssrn.5528763.

[8] G. Maddali, “An Efficient Bio-Inspired Optimization Framework for Scalable Task Scheduling in Cloud Computing Environments,” *Int. J. Curr. Eng. Technol.*, vol. 15, no. 3, pp. 229–238, 2025.

[9] E. Esteves, G. Locatelli, N. A. Bou, and R. S. Ferrarezi, “Sap Analysis: A Powerful Tool for Monitoring Plant Nutrition,” *Horticulturae*, vol. 7, no. 11, 2021, doi: 10.3390/horticulturae7110426.

[10] N. P. Bhat, “Optimizing Scrap Management in Production Planning Processes,” *Int. J. Comput. Trends Technol.*, vol. 71, no. 5, pp. 48–52, 2023, doi: 10.14445/22312803/ijctt-v71i5p108.

[11] J. E. Kofi, “Data-Driven Cloud Workload Optimization Using Machine Learning Modeling for Proactive Resource Management,” *Int. J. Emerg. Res. Eng. Technol.*, vol. 6, no. 4, pp. 27–37, 2025, doi: 10.63282/3050-922X.IJERET-V6I4P104.

[12] S. Kumara, “Identity-Driven IoT Security in Telecom Ecosystems: Implications for Scalable and Trustworthy Digital Infrastructure,” *Int. J. Appl. Math.*, vol. 38, no. 12s, pp. 2797–2816, 2025.

[13] P. Samaranayake and T. Laosirihongthong, “Production planning and scheduling using integrated data structures in ERP: Implementation and numerical simulation,” *Int. J. Manag. Sci. Eng. Manag.*, vol. 10, no. 3, pp. 176–190, 2015, doi: 10.1080/17509653.2014.939726.

[14] V. K. C. Mandavilli, “The Transformative Power of SAP AI Across Industries: A Technical Overview,” *J. Comput. Sci. Technol. Stud.*, vol. 7, no. 4, pp. 358–367, May 2025, doi: 10.32996/jcsts.2025.7.4.42.

[15] S. Amrale, “Proactive Resource Utilization Prediction for Scalable Cloud Systems with Machine Learning,” *Int. J. Res. Anal. Rev.*, vol. 10, no. 4, pp. 758–764, 2023.

[16] R. Patel, “Survey of Digital Twin Applications in Predictive Maintenance for Industrial Material Handling Equipment,” *Journal, Int. Recent. O F Sci. Technol.*, vol. 9, no. 4, pp. 77–86, 2024.

[17] D. Gadve and G. Deshmukh, “SAP Production Planning (PP) in Manufacturing: A Systematic Review of Digital Transformation and Integrated Business Functions for Sustainable Growth,” *Int. J. Sci. Res. Eng. Manag.*, vol. 10, pp. 1–9, 2026, doi: 10.55041/IJSREM.IBFE148.

[18] K. Immidi and S. Mane, “Integration of Artificial Intelligence with SAP Integrated Business Planning (IBP): A Technical Analysis of Implementation Success and Performance Metrics Across Industries,” *Int. J. Eng. Res. Technol.*, vol. 14, no. 2, pp. 1–11, 2025.

[19] T. Vaidya, “Digital Twin-Driven Production Planning in SAP S/4HANA: A Case for Predictive and Adaptive Supply Chains,” *J. Comput. Sci. Technol. Stud.*, vol. 7, 2025, doi: 10.32996/jcsts.2025.7.7.30.

[20] Q. Fu, “Implementation Standards and Industry-Specific Adaptation of SAP-Lean Production Integration in Manufacturing: A Multi-CasValidation with Quantitative Performance Optimization,” *Front. Manag. Sci.*, vol. 4, no. 5, pp. 67–74, Oct. 2025, doi: 10.63593/FMS.2788-8592.2025.09.008.

[21] K. Chenna, “Hyper Automation in SAP Production Planning,” *Int. J. Comput. Eng. Technol.*, vol. 15, no. 5, pp. 505–513, 2024, doi: 10.5281/zenodo.13850946.

[22] A. H. Bachchan, “AI and ML in SAP Manufacturing Supply Chains: Enabling Predictive Quality and Process Control,” *Int. J. Res. Appl. Innov.*, vol. 6, no. 6, 2023.

[23] M. Kohli, “Using Machine Learning Algorithms on data residing in SAP ERP Application to predict equipment failures,” *Int. J. Eng. Technol.*, vol. 7, pp. 312–319, 2017.

[24] M. Kohli, “Supplier Evaluation Model on SAP ERP Application using Machine Learning Algorithms,” *Int. J. Eng. Technol.*, pp. 306–311, 2017.

[25] A. Papacharalampopoulos, K. Alexopoulos, P. Catti, P. Stavropoulos, and G. Chryssolouris, “Learning More with Less Data in Manufacturing: The Case of Turning Tool Wear Assessment through Active and Transfer Learning,” *Processes*, vol. 12, no. 6, 2024, doi: 10.3390/pr12061262.

[26] M. Kalal, "Lean-Driven SAP Production Planning Optimization Using Machine Learning for Inventory and Throughput Efficiency," 2026 14th International Symposium on Digital Forensics and Security (ISDFS), Boston,

MA, USA, 2026, pp. 1-6, doi: 10.1109/ISDFS69419.2026.11459029. keywords: {Radio frequency;Protocols;Radio access networks;Regional area networks;HTTP;Communication systems;Radio communication;Synchronous digital hierarchy;Telecommunications;Plugs;SAP;demand planning;predictive analytics;inventory management;enterprise resource planning;machine learning},

[27] Kalal, M. (2026). SUPPLY CHAIN RESILIENCE THROUGH INTEGRATED SAP PRODUCTION PLANNING–PROCUREMENT SYNCHRONIZATION. *Power System Protection and Control*, *54*(1), 620-637.

[28] M. Kalal, "Predictive Production Planning in Advanced Manufacturing Using SAP PP and Analytics," 2026 5th International Conference on Sentiment Analysis and Deep Learning (ICSADL), Birendranagar, Nepal, 2026, pp. 1363-1370, doi: 10.1109/ICSADL67539.2026.11452022. keywords: {Production planning;Decision making;Production;Predictive models;Manufacturing;Fourth Industrial Revolution;Resource management;Internet of Things;Predictive analytics;Optimization;Predictive Production Planning;SAP PP;Advanced Manufacturing;Analytics;Industry 4.0;Supply Chain Optimization},

[29] M. Kalal, "Integrating SAP PP and SAP Digital Manufacturing for Lean Production Optimization," 2026 International Symposium of Systems, Advanced Technologies and Knowledge (ISSATK), Hammamet, Tunisia, 2026, pp. 1-7, doi: 10.1109/ISSATK69667.2026.11566800. keywords: {Modeling;Manufacturing;Optimization;Printing;Machine learning;Accuracy;Artificial intelligence;Production planning;Timing;Machining;Digital manufacturing;(Systems, Applications, and Products) SAP;Production planning (PP);Lean Production Optimization;Machine Learning},