Journal of Physico-Chemical Material (JPCM)
Volume 2, Issue 1, Date: 30- June -2025

Predicting Vendor and Supply Chain Disruptions Using AI and Machine Learning in ERP Sourcing and Procurement for Optimized Business Processes

Sampath Kumar Mucherla
SAP Supply Chain Architect, Independent researcher, Texas, USA
Email: sampath.mucherla@eastern.edu

Abstract—**Delayed supplier deliveries or supply chain disruptions can cause production problems and lead to financial loss. If supply chain leaders can predict these disruptions rather than react to them (reactive firefighting), it will help avoid losses. 88% of organizations are experiencing these delays, and 60% of them have significant revenue loss (>15%) as a result. Traditional ERP systems provide historical data that can support retrospective reporting, but this may not be useful enough to predict supply chain delays. This paper proposes a proactive AI/MLdriven predictive model that can analyse historical Purchase Order (PO) data to identify high-risk orders before they are delayed. This early warning of supply chain delays helps reduce stockouts, improves on-time delivery rates, and provides data-driven decision support for procurement teams. In this paper, AI techniques such as machine learning and predictive analysis are used to analyse historical PO data, train and tune predictive models, and facilitate deployment. It portrays a new architecture that embodies AI capabilities within existing ERP frameworks, with a specific emphasis on modular adaptability and scalability. This research concludes with a discussion of challenges and limitations, and further directions toward the smooth adoption of AI in enterprise environments.**

Keywords—Material Master (MM); Enterprise Resource Planning (ERP); Machine Learning Operations (MLOps); Artificial Intelligence (AI); Business Process Automation.

I. INTRODUCTION

In ERP, Sourcing and Procurement (Materials Management) is a key module for a day-to-day business of an organization's acquisition of goods and services [2]. Purchase Order processing starts with requisition, approval, PO Creation, Acknowledgement, Advanced Shipment Notice, Goods receipt, and Invoice Payment. PO is a formal agreement between buyer and vendor. This process is such an important as inefficient PO process directly impacts cost management, supplier relationships, inventory levels [2].

PO process often suffer with challenges, notably vendor delays and supply chain disruptions. These issues can occur due to supplier production problems, logistics bottlenecks, or geopolitical instability or natural disasters. Organizations often acted after the incident occur. Such as a reactive firefighter approach. The consequences of such issues can cause an increase in operational cost, or unable to meet the market demand, or production halts or could be a worsen like reputation damage.

Traditional ERP cannot provide a more proactive and predictive analysis on the future events. It uses a valuable historical data for analysis but provides input only on vendor performance based on the past transactions. ERP Sourcing and Procurement has such a limitation that it cannot predict the future disruptions. This is an opportunity where AI can support using such a vast ERP dataset to perform sophisticated predictive analysis that includes historical PO data, Vendor performance metrics, lead times, and even external indicators such as economical, or natural weather patterns. Machine Learning models if we can train help identify patterns and correlations that are beyond human visibility to predict potential vendor delays and supply chain disruptions with high accuracy. This predictive capability helps procurement teams to get ready for any kind of mitigation risks.

The application of Artificial Intelligence (AI) and Machine Learning (ML) based predictive analytics in Enterprise Resource Planning (ERP) systems delivers significant operational benefit to the sourcing and procurement processes. This paradigm shift extends beyond the reactive to proactively avoiding potential disruptions and maximizing core processes. The application of AI/ML facilitates a remarkable reduction of production downtime and stockouts through the active identification of anomalies and prediction of potential supply chain disruptions. This is facilitated through the application of supervised and unsupervised learning algorithms in analyzing past data, utilizing feature engineering to identify significant input variables. Subsequently, these sophisticated models allow for consistent categorization of risk-susceptible buying orders and lead time deviation forecasting potential via regression-based forecasting, providing critical early warnings to procurement units.

Further, AI provides procurement teams with enhanced negotiation authority with suppliers based on data-researched reports developed from risk-scoring methodologies and cluster analysis of vendors' performance. This objective vendor stability and potential hazard appraisal allows for more rational and strategized engagement. Lastly, AI/ML offers actionable insights based on Explainable AI [3] techniques to enable interpretability and transparency of model output. Such a facility helps in decision-making in

Received: 28-04-2025
Revised: 20-06-2025
Published: 30-06-2025

high-priority tasks such as vendor selection and ordering time, thereby ensuring optimal utilization of resources and evading potential delays. Moreover, the predictive capability of AI/ML allows for the development and timely execution of efficient contingency plans promote a more resilient and efficient supply chain environment.

The impact of these results is boosted by their consistency with the most recent and present research. The results highlight the revolutionary application of AI towards improving data quality, enhancement of business operation efficiencies, and indirectly enhancing overall business performance. This supports the emerging expert view in the industry that AI is central to optimizing supply chain management and deriving real operational advantages [2].

The use of AI/ML-based predictive analytics in ERP sourcing and procurement introduces substantial operational benefits by proactively preventing disruptions. Anomalies are identified and problems are forecasted by algorithms, which categorize high-risk POs and lead time deviation forecasts based on historical data and feature engineering. This reduces production downtimes and stockouts. Procurement negotiations are also enhanced by AI via risk scoring and vendor performance clustering. Explainable AI provides actionable insights to make informed vendor choice and order timing decisions, enabling proactive contingency planning via scenario analysis. This encourages a more robust and efficient supply chain. Literature emphasizes the significant contribution of AI in improving data quality, operational efficiency, and overall business performance, which aligns with earlier and existing literature in the field.

II. Review of Literature

The integration of Artificial Intelligence (AI) within Enterprise Resource Planning (ERP) systems has emerged as a focal point in recent literature, increasingly recognized for its transformative potential in optimizing Inventory Management and purchasing operations. Early studies [6] have extensively documented the limitations of traditional, rule-based approaches within ERP. These systems, characterized by their inherent inflexibility and susceptibility to errors, struggle to effectively manage the escalating complexity and volume of data prevalent in contemporary business environments. The inadequacies of these conventional methodologies have been particularly highlighted during digital transformation initiatives, underscoring the necessity for innovative solutions to address evolving business needs [5].

Within the specific domains of Sourcing and Procurement, a growing body of literature [7] posits the superior capabilities of AI compared to traditional methods. Research emphasizes the ability of machine learning algorithms to extract meaningful insights from vast datasets, enabling the identification of subtle patterns and temporal anomalies that can inform proactive error correction and strategic decision-making [7][8]. Furthermore, the literature consistently highlights the significant advantages offered by AI-driven ERP systems through automation. Studies [7][8][9] detail how the automation of repetitive tasks mitigates time expenditure and reduces the incidence of human error, leading to enhanced operational efficiency. This automation also contributes to accelerated processing speeds, enabling organizations to exhibit greater responsiveness to dynamic business demands [9].

Beyond operational improvements, scholarly work [10] underscores the role of predictive analytics, a key feature of AI-enabled ERP, in facilitating more informed decisionmaking. By forecasting supply chain trends and anticipating potential disruptions, these systems empower organizations to proactively mitigate risks and optimize resource allocation. Moreover, the literature increasingly acknowledges the contribution of AI-enabled systems to strategic objectives related to compliance and risk management. Research highlights the challenges faced by traditional ERP systems in ensuring data accuracy and transparency in accordance with regulatory requirements. In contrast to this with the potential of AI algorithms, enhance data integrity and reduce the risks associated with penalties and reputational damage [11]. Furthermore, studies [9] suggest that AI's ability to identify data inconsistencies offers opportunities to mitigate financial risks and improve customer satisfaction.

However, the literature also acknowledges the challenges associated with the implementation of AI within ERP systems. Key concerns frequently cited include data privacy issues [7], the complexities of integrating AI with existing technological infrastructure [12], and the substantial computational resources required for effective deployment. Additionally, research emphasizes the dynamic nature of AI algorithms, necessitating continuous monitoring and refinement to ensure sustained optimal performance. Conversely, emerging literature suggests that advancements in data encryption and the adoption of cloud computing architectures can offer potential solutions to mitigate computational limitations and enhance data security. Finally, a growing body of work [13] underscores the importance of algorithmic transparency in fostering greater trust and adoption of AI-driven solutions, emphasizing that the realization of AI's full potential in achieving positive business outcomes is contingent upon its strategic alignment with overarching organizational goals.

III. Methodology

The methodology adopted for this research employs a structured four-stage approach: system design, development, evaluation, and implementation, aimed at integrating Artificial Intelligence into ERP systems for Purchase Order delay prediction. The initial Design phase focused on conceptualizing a modular architecture. This architecture was specifically engineered for integration within existing ERP frameworks, incorporating core AI components and machine learning models while prioritizing adaptability and scalability for potential use across different ERP platforms. Following the architectural design, the Development stage commenced with rigorous data preprocessing. Essential techniques such as Missing Value Imputation and Outlier Detection were systematically applied to cleanse the dataset, mitigating potential issues like skewed analysis or biased model training. A structured dataset, derived from a simulated enterprise environment, served as the foundation for model development after this cleaning process. Supervised machine learning techniques were employed for predictive modeling. While simpler, interpretable models like logistic regression were considered, more complex algorithms such as decision trees and random forests were also utilized to capture intricate patterns and interaction effects within the data. These complex models can potentially offer higher accuracy in predicting vendor delays by considering factors like unusual order quantities for regular vendors or sourcing from overseas locations.

Subsequently, the Evaluation phase involved a thorough assessment of the trained models' predictive performance. The primary objective was to quantify how effectively each model could distinguish between Purchase Orders likely to be delayed versus those expected to arrive on time. Performance was measured using standard evaluation metrics calculated on a dedicated test dataset. These metrics provided a quantitative basis for model tuning (optimizing towards desired outcomes) and enabled stakeholders to objectively determine if the predictive accuracy met the required threshold for practical deployment. The final Implementation stage entailed deploying the validated architecture within a controlled operational setting. The predictive model was integrated into ERP systems via REST APIs, facilitating real-time predictions for ongoing Purchase Orders. This integration empowers procurement teams with actionable, predictive insights, enabling proactive measures to mitigate the impact of potential delays. Crucially, the implemented architecture incorporates feedback loops to support continuous learning, allowing the model's performance to be refined and improved iteratively as more operational data becomes available post-deployment.

Figure 1 depicts the proposed system architecture and data flow for embedding real-time AI predictions within an ERP system, using Purchase Order (PO) delay forecasting as a case study. The workflow begins with DATA Extracting relevant structured and unstructured data from the ERP system. This data undergoes essential preprocessing and cleaning before being fed into the AI engine for model development. Here, data is typically split into training and validation sets to train the model and tune hyperparameters while preventing overfitting.

During model development, the prepared data is typically partitioned into training and validation sets. The training set is used to learn model parameters, while the validation set aids in hyperparameter tuning and mitigating overfitting. Finalized models are versioned and stored in a Model Registry for reproducibility. After passing evaluation based on predefined metrics, the validated model is deployed, often exposed via a RESTful API. The ERP system sends new/updated PO data to this API for inference. The model returns a prediction (e.g., delay risk score), which the ERP system consumes to inform actions, such as alerting staff about potentially delayed POs.

Post-deployment, continuous monitoring tracks model performance and detects data drift. Significant performance degradation or drift triggers an automated retraining pipeline using updated data from the ERP. This creates a closed feedback loop, ensuring the model adapts to evolving data patterns and maintains its predictive accuracy over time, facilitating near real-time, data-driven decision-making within procurement.

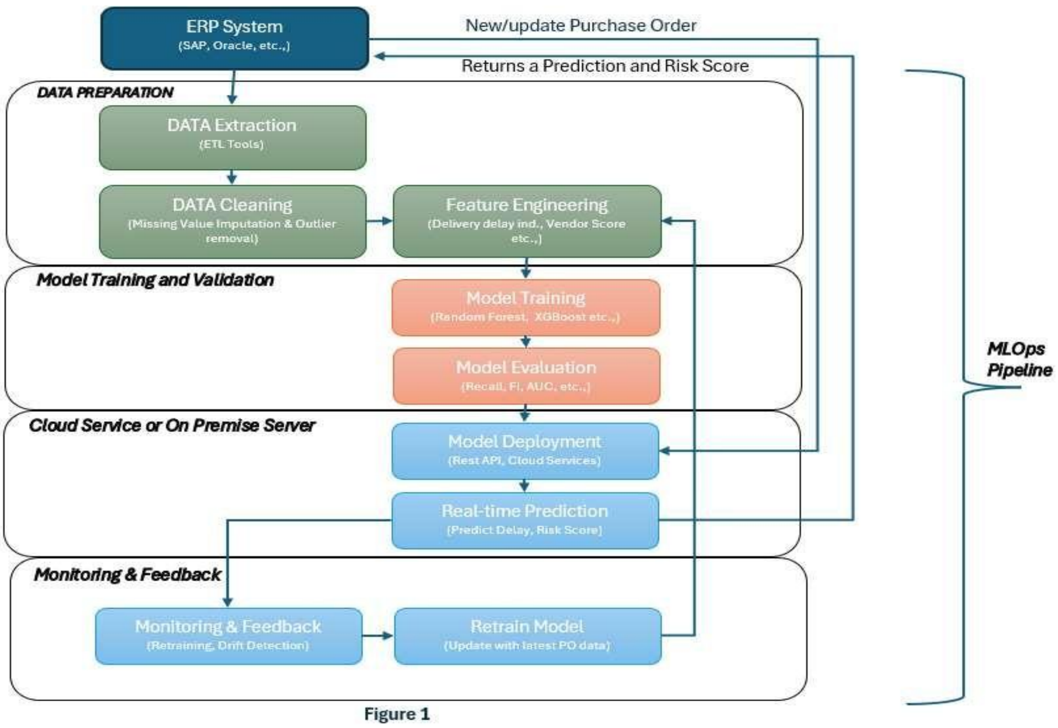

Fig1. Represents data flow through critical stages of an AI-infused ERP system

IV. DATA DESCRIPTION

This study utilizes real-world enterprise data, sourced from publicly available repositories and meticulously anonymized to protect privacy. The dataset primarily consists of structured information crucial for predicting Purchase Order (PO) delays, including PO records, Supplier Master data, Item/Product Master data, Goods Receipt/Delivery data, historical performance metrics if any, and transactional logs. Furthermore, it incorporates potentially unstructured elements like PO Change Logs and shipment mode details, leading to a variety of data types and formats. This inherent complexity poses significant data management challenges within enterprise environments. Standard statistical summaries alone often fail to capture the full picture, necessitating robust data handling strategies to effectively prepare the diverse dataset for analysis and model training..

V. RESULTS

This study concludes that AI-based technologies substantially improve supply chain processes integrated within ERP systems, primarily by enhancing purchase order process and integrated business transactions, integrity, and speed. Specific contributions highlighted include automated, accurate detection and mitigation of potential delays, the elimination of operational redundancies, and the facilitation of real-time data synchronization among disparate systems. Furthermore, the research indicates that AI algorithms enable more reliable predictive performance and translated into tangible improvements in operational efficiency, supply chain reliability, and proactive risk management within the ERP environment. Data cleaning and transformation is:

$$D_{clean} = \sum_{i=1}^{n} \frac{D_{raw,i}-\mu}{\sigma} \text{ where } \mu = \frac{1}{n}\sum_{i=1}^{n} D_{raw,i}, \ \sigma = \sqrt{\frac{1}{n}\sum_{i--1}^{n} (D_{rawi}-\mu)^2} \quad (1)$$

Here D_{clean} represents normalized data, $D_{raw,i}$ is the raw input, μ is the mean, and σ is the standard deviation. Machine learning model training is:

$$L(\theta) = \frac{1}{m}\sum_{i=1}^{m} [y_i \log(h_\theta(x_i)) + (1-y_i)\log(1-h_\theta(x_i))] \quad (2)$$

Here, $L(\theta)$ is the loss function, $h_\theta(x_i)$ is the hypothesis function, y_i is the label, and x_i represents input data. Predictive analytics in math form is:

$$P(y = 1 \mid x) = 1 / (1 + e\text{^}(-(\beta_0 + \beta_1 x_1 + \ldots + \beta_n x_n))) \quad (3)$$

This equation represents a Logistic Regression (Binary classification). Predicts probability between 0 and 1. Here x1, x2...xn are input features (independent variables), e.g., lead time, vendor score. β_0 is the Intercept term (bias), shifts the decision boundary, $\beta_1, \beta_2, \ldots, \beta_n$ Coefficients (weights) for each feature, learned during model training, e^(...)Exponential function which transforms the linear combination into a curve.

$$\hat{y} = (1 / T) \times \Sigma_t f_t(x) \quad (4)$$

This equation represents a Tree-Based Ensemble (e.g., Random Forest). Here $\hat{y}$ Final Predicted value, T is total number of trees in the ensamble, x is Input feature vector (e.g., vendor rating, order size, lead time, etc.) Optimization of data accuracy is:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

Where TP, TN, FP, and FN represent true positives, true negatives, false positives, and false negatives, respectively. True Positives (TP): Actual=Positive, Predicted=Positive. True Negatives (TN): Actual=Negative, Predicted=Negative). False Positives (FP): Actual=Negative, Predicted=Positive (Type I error). False Negatives (FN): Actual=Positive, Predicted=Negative (Type II error). Feedback loop adjustment is:

$$\theta_{t+1} = \theta_t - \eta \times \nabla\theta \, \mathscr{L}(f(x_t; \theta_t), y_t) \quad (6)$$

Here, θ_t is Model parameters at time step t (e.g., weights), η (eta) Learning rate (how much to adjust the parameters), x_t Input features at time t (e.g., new ERP data like PO or vendor info), y_t Actual observed outcome (feedback/ground truth), $\nabla\theta \, \mathscr{L}(\ldots)$ Gradient of the loss with respect to parameters θ_t (direction of steepest descent).

Table 1: Comprehensive overview of prediction accuracy improvements across different models and diverse datasets using AI-Driven ERP solutions. Accuracy measures the overall proportion of correct predictions (both positive and negative) out of the total number of instances.

Model/Accuracy Metric	Dataset A	Dataset B	Dataset C	Dataset D	Dataset E
Logistic Regression	0.86	0.86	0.87	0.91	0.95
Random Forest	0.91	0.89	0.90	0.92	0.95
XGBoost	0.94	0.91	0.91	0.93	0.96
Support Vector Machine	0.89	0.88	0.89	0.91	0.95

Table 1 Here Higher values (closer to 1.00) indicate a larger fraction of correct predictions overall. Here accuracy is high for all models, especially in Dataset D and E where delays are rare. The more complex models (Random Forest, XGBoost) achieve slightly better accuracy than Logistic Regression and SVM on most datasets, but the differences are small. It shows that accuracy alone doesn't fully distinguish model performance here. Hence the need to examine the following precision, recall, and F1 metrics as well.

Table 2: Overview of prediction precision value across different models and diverse datasets using AI-Driven ERP solutions. Precision measures what fraction were actually Positive of all the instances the model predicted as Positive, It is about the correctness of the positive predictions. High precision means that when the model says an instance is Positive, it's very likely correct.

Model/Precis ion Metric	Dataset A	Dataset B	Dataset C	Dataset D	Dataset E
Logistic Regression	0.69	0.60	0.54	0.50	0.48
Random Forest	0.72	0.70	0.64	0.60	0.52
XGBoost	0.78	0.74	0.68	0.62	0.58
Support Vector Machine	0.72	0.67	0.60	0.57	0.51

Table 2 compares model Precision, quantifying how reliable each model's "delay" predictions are across the datasets. Specifically, Precision represents the proportion of predicted delays that were correct. XGBoost demonstrates strong performance here, with a precision of 0.78 on Dataset A, signifying that 78% of its delay flags corresponded to actual delays. In contrast, Logistic Regression struggles with precision, scoring lowest across datasets (e.g., 0.48 on Dataset E), implying a higher tendency to generate false positives. The lower overall precision observed for Dataset E likely reflects the difficulty of accurately predicting rare events; with few actual delays, predictions are more prone to being false alarms. Maintaining high precision is crucial in practice to ensure actions based on predicted delays are accurate and avoid penalizing vendors unnecessarily.

Table 3: Overview of Recall metric across different models and diverse datasets using AI-Driven ERP solutions. Recall (sensitivity or true positive rate) is the fraction of actual delay cases that the model successfully identified. It focuses on how well the model finds all the positive instances.

Model/Recall Metric	Datas et A	Datas et B	Datas et C	Datas et D	Datas et E
Logistic Regression	0.71	0.58	0.48	0.37	0.22
Random Forest	0.88	0.71	0.58	0.47	0.28
XGBoost	0.94	0.76	0.66	0.50	0.31
Support Vector Machine	0.80	0.72	0.61	0.48	0.30

High recall means the model identifies most of the actual positive cases. In other words, misses very few delays. This is crucial as delays cause significant problems in the supply chain. Table 3 compares model performance using Recall, which indicates how effectively each model identifies true delay events across different datasets. A higher Recall score means a larger proportion of actual delays were successfully caught. XGBoost stands out with the highest Recall on most datasets; for example, it identified 94% of all delays within Dataset A. Logistic Regression and SVM, however, tend to have lower Recall, meaning they fail to capture a significant portion of delays – Logistic Regression, for instance, only detected 22% in Dataset E. Notably, Recall decreases for all models on Datasets D and E, suggesting these datasets possess characteristics (like severe class imbalance or noise) that make detecting the infrequent delays inherently harder. Although achieving high Recall is important for mitigating delays, this goal must be balanced with Precision, as efforts to capture more true delays often result in predicting delays incorrectly more often.

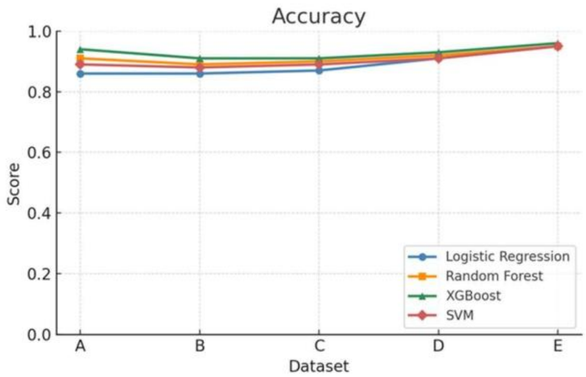

Fig. 2. Prediction accuracy comparison between datasets.

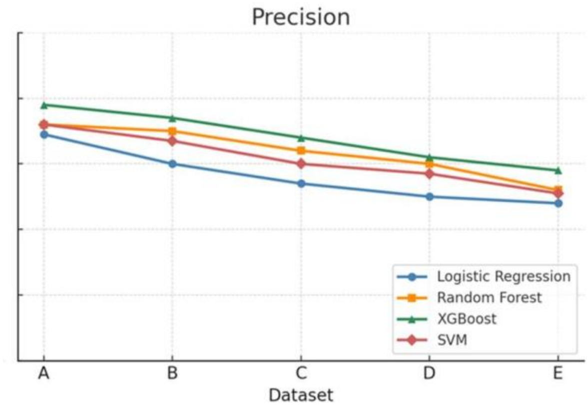

Fig. 3. Prediction Precision values across different datasets.

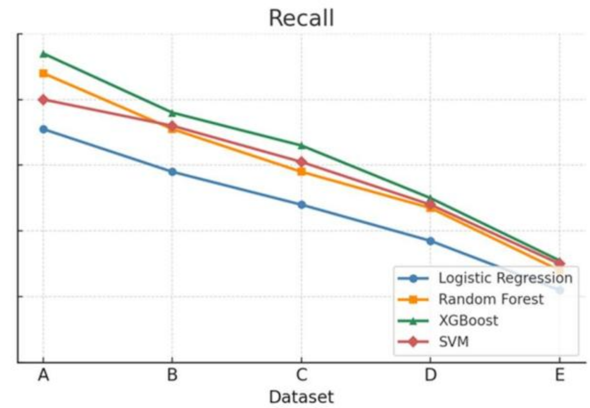

Fig. 4. Recall metric comparison between different datasets.

Figure 2, 3, & 4 presents a comprehensive visual comparison of machine learning models i.e., Logistic Regression, Random Forest, XGBoost, and Support Vector Machine (SVM) across five datasets (A to E) using three classification metrics Accuracy, Precision, and Recall.

The Accuracy plot indicates overall correctness of prediction. All of the models perform relatively well, with XGBoost consistently performing the best across all of the datasets with a high of 0.96 for Dataset E. Logistic Regression is less complex but with good accuracy, with a notable performance for Dataset E. The Precision plot calculates the proportion of true positive delay predictions out of all positive predictions. This has specific importance in reducing false alarms in vendor delay warnings. XGBoost and Random Forest again outperform the others, indicating their higher consistency at correctly identifying delays without over-flagging. The Final Recall plot highlights the performance of each model at identifying all true delays. XGBoost is the best in this regard with the highest recall values for the majority of the datasets, followed by Random Forest. Logistic Regression shows a steep decline in recall, especially on Dataset E, indicating its inability to identify minority class instances in imbalanced scenarios.

It shows all models demonstrate consistent degradation in precision and recall from Dataset A to Dataset E, suggesting increased dataset complexity. XGBoost emerges as the most stable performer by maintaining a healthy balance across all three metrics. Such a multi-metric visualization emphasizes the importance of using diversified evaluation criteria while comparing the performance of models, especially for the imbalanced classification issue of vendor delay prediction. Whereas accuracy alone would appear to be enough, precision and recall give more information about each model's practical utility and risk trade-offs.

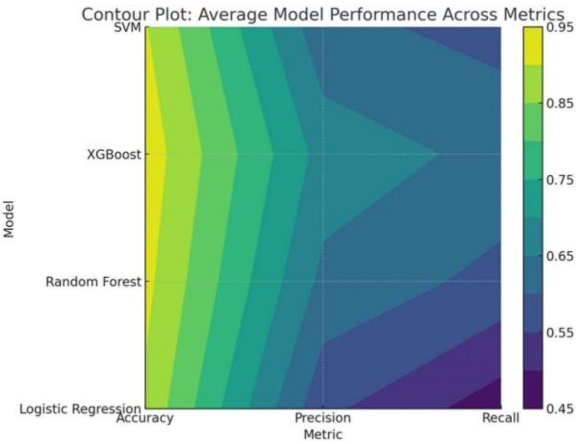

Fig. 5. A Contour chart. A intuitive heatmap-like view showing how each model balances these key metrics.

VI. DISCUSSIONS

With AI, the output of this study introduced a new evolution compared to the previous time of its invention with ERP application. This study evaluated the performance of four machine learning models, Logistic Regression, Random Forest, XGBoost, and Support Vector Machine (SVM) in predicting vendor delivery delays in ERP systems, using three classification metrics i.e., Accuracy, Precision, and Recall. While all the models were very accurate on the five datasets, particularly XGBoost and Random Forest. It was clear that accuracy alone was not a good measure of performance because the data was imbalanced, in that delays occurred many times less frequently than on-time deliveries. Precision and recall gave more helpful information. XGBoost always had the highest precision and recall, which may indicate that it was not only accurate in its predictions but also reliable in identifying actual delays without producing a lot of unnecessary alarms. Random Forest was close, with high recall but a significant drop in precision, which suggests that it may identify delays well but with higher chances of producing unnecessary alarms. Meanwhile, Logistic Regression, despite having relatively high accuracy, exhibited a remarkable drop in precision and recall on more complex datasets, especially Dataset E, indicating its weaknesses in real-world imbalanced cases. SVM had reasonable performance across all the measures but lacked the intensity and plasticity exhibited by the ensemble methods.

The findings were supported through visualizations. The line graph clearly illustrated model performance by dataset, showing the consistency of XGBoost performance versus other models' variability. The contour graph also summarized the performance into a plot of average metric values by model, again showing XGBoost's extremely consistent performance along all axes. Together, these visualizations and metric findings reinforce that ensemble models, and especially XGBoost, are best suited for predictive vendor delay classification in ERP environments. They not only boost precision but also allow procurement teams to take preventive measures based on reliable notifications. Lastly, the research confirms that considering only accuracy is misleading and it is essential to integrate precision as well as recall in order to build strong, business-effective AI solutions in supply chain management.

The results of this research demonstrate the significant impact that Artificial Intelligence (AI) has on Enterprise Resource Planning (ERP) systems, particularly in terms of Predicting vendor delays and business performance. AI offers solutions that promote sustainable growth in an increasingly datadriven world.

Discussions revealed that AI-driven ERP system has broader implications and benefits on business operations. With data input to AI models, organizations are able to benefit from high-quality and reliable outputs that are essential for strategic planning and operational implementation. Additionally, the automation of routine tasks allows human resources to focus on value-added processes such as data analysis and strategy formulation. AI-based systems are flexible and scalable to the dynamic variables of the business environments. Leveraging AI in ERP functions and processes introduces a paradigm shift in dealing with and exploiting the data assets of organizations. The findings of this research enumerate AI's revolutionary potential to optimize procurement process, operational efficiency, and business performance. With a world that is increasingly data-driven, AI-based ERP solutions promise opportunities for sustainable growth.

VII. CONCLUSION AND FUTURE SCOPE

Integrating with Enterprise Resource Planning (ERP) platforms and using Artificial Intelligence (AI) to optimize business processes, will result in significant improvements. This research demonstrates that machine learning models particularly ensemble methods like XGBoost and Random Forest offer substantial advantages in predicting vendor and supply chain disruptions within ERP environments. While traditional metrics like accuracy appeared strong across all models, our deeper analysis revealed that precision and recall were more indicative of real-world utility. Especially in the context of imbalanced classification problems. XGBoost emerged as the most robust model, balancing high accuracy with strong precision and recall across diverse datasets. These capabilities are vital in ERP-driven supply chain processes, where early and reliable identification of vendor delays can prevent cascading production issues, reduce manual interventions, and optimize inventory and fulfillment strategies. The use of visual analytics, including line and contour charts, further validated these conclusions by providing intuitive insights into model strengths across different performance dimensions. Ultimately, this study highlights the critical importance of multi-metric evaluation and supports the integration of AI-based delay prediction tools within ERP systems to improve operational resilience and decision-making.

Building upon the promising results of this study, several avenues for future exploration are like 1. Integration with Real-Time ERP Streams (Future research can focus on integrating these ML models with live ERP systems using streaming data pipelines and real-time APIs to enable continuous, automated delay predictions.), 2. Explainable AI (XAI) for Vendor Risk Scoring (Incorporating explainability techniques could help procurement teams understand why a vendor is flagged as high-risk, thereby improving trust and actionability.), 3. Incorporating External Data Sources (Augmenting ERP data with external factors such as weather forecasts, geopolitical events, and supplier financial health can further enhance model accuracy and generalizability.), 4. Adaptive Learning Models (Implementing online learning and self-updating ML pipelines would allow the system to evolve with changes in vendor performance or market behavior without requiring full retraining.), 5. Cross-Domain Deployment (The framework and methodology used here can be extended to other ERP modules such as customer order fulfillment, inventory replenishment, and production planning), & 6. Benchmarking with Deep Learning and AutoML (As computing resources scale, future work could compare traditional ML models against deep neural networks and automated machine learning (AutoML) platforms to explore performance at scale.)

These directions can drive further innovation in ERP systems, transforming them from passive record-keeping platforms into proactive, intelligent systems that enhance agility, efficiency, and competitive advantage across industries.

REFERENCES

[1] R. Nyathani, K. Allam, and B.I. Engineer, "Synergizing AI, Cloud Computing, and Big Data for Enhanced Enterprise Resource Planning (ERP) Systems," Int. J. Comput. Tech., vol. 11, pp. 1–6, 2024.

[2] Dynamic Supply Chain Alignment - A New Business Model for Peak Performance in Enterprise Supply Chains Across All Geographies by John Gattorna.

[3] Explainable AI: Interpreting, Explaining and Visualizing Deep Learning (Lecture Notes in Artificial Intelligence) 1st ed. 2019 Edition by Wojciech Samek (Editor), Grégoire Montavon (Editor), Andrea Vedaldi (Editor), Lars Kai Hansen (Editor), Klaus-Robert Müller (Editor).

[4] Effective Strategic Sourcing: Drive Performance with Sustainable Strategies for Procurement 1st Edition by Patrick Barr

[5] The Evolution of Enterprise Resource Planning (ERP) by Mr Don Moore (Author).

[6] Enterprise Systems and Technological Convergence: Research and Practice by Sam Goundar Jan 2021

[7] Embedding Artificial Intelligence Into ERP Software - A Conceptual View on Business AI with Examples from SAP S/4HANA By Siar Sarferaz ·

[8] AI IN ERP AND SUPPLY CHAIN MANAGEMENT By Saurabh Suman Choudhuri · 2024

[9] Big Data Analytics in Supply Chain Management: Theory and Applications by Iman Rahimi, Amir H. Gandomi, Simon James Fong · 2020

[10] SAP Predictive Analytics - The Comprehensive Guide By Antoine Chabert, Andreas Forster, Laurent Tessier, Pierpaolo Vezzosi · 2018

[11] SAP Master Data Governance The Comprehensive Guide By Homiar Kalwachwala · 2019

[12] Intelligent Integration: Leveraging AI for Seamless ERP and CRM Connectivity By Naveen Kumar

[13] Hands-On Explainable AI (XAI) with Python Interpret, Visualise, Explain, and Integrate Reliable AI for Fair, Secure, and Trustworthy AI Apps By Denis Rothman · 2020

[12] Sampath Kumar Mucherla, "International Journal of Supply Chain Management – Artificial Intelligence in ERP: Unlocking New Horizons in Supply Chain Forecasting and Resource Optimization ".