Journal of Information Systems Engineering and Management
2025, 10(52s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Optimizing Procurement Systems Using Classification-Based Supply Delay Predictions

Kasthalani[1] Muhammad Zarlis[2*]
[1,2] Master of Information Systems Management, Bina Nusantara University, Jakarta, Indonesia 11480
Email: [1]kasthalani@binus.ac.id, [2]muhammad.zarlis@binus.ac.id

ARTICLEINFO

Received: 17 Dec 2024

Revised: 15 Feb 2025

Accepted: 26 Feb 2025

ABSTRACT

Supply delays in procurement can disrupt production operations and impact supply chain efficiency. This study applies data mining classification techniques to predict supply delays in a state-owned fertilizer company in Indonesia. Using the CRISP-DM methodology, we evaluated Logistic Regression, Support Vector Machine (SVM), Random Forest, and Deep Learning models. The results indicate that Deep Learning achieved the highest AUC, demonstrating superior predictive capability, while Random Forest also performed well. Key factors influencing supply delays include order value and Incoterms. These findings can support the development of an early warning system to enhance procurement efficiency and mitigate supply risks.

Keywords: Predictive Analytic, Machine Learning, Data Mining, Classification, Supply Chain Management

I. INTRODUCTION

Rocurement of goods is an important stage in upstream supply chain management in an organization or company to ensure smooth operations. An efficient procurement process directly impacts the overall supply chain efficiency by reducing lead time, ensuring timely stock availability, and optimizing acquisition costs [1]. In this context, information systems serve as the backbone of supply chain management, enabling real-time data collection and exchange across different supply chain entities, including suppliers, manufacturers, distributors, and consumers [2]By leveraging accurate and timely information, companies can make informed decisions regarding production planning, inventory management, and distribution. The integration of enterprise resource planning (ERP) systems and e-procurement platforms has significantly enhanced procurement transparency and efficiency by automating ordering, shipping, and payment processes [3].

However, despite the implementation of digital procurement solutions, supply delays remain a persistent challenge in procurement operations. At PT XYZ, a subsidiary of a state-owned fertilizer producer in Indonesia, an audit by the Supreme Audit Agency identified supply delays in approximately 28% of purchase orders issued between 2022 and the first half of 2024. Timeliness is a critical factor in procurement, which emphasize six key principles: right quantity, right quality, right price, right supplier, right location, and right time [4]. Delays in procurement can disrupt operational performance and cause financial losses. Given the need for proactive measures to address supply delays, advanced analytical approaches such as data mining and machine learning are being explored to improve supply accuracy and mitigate risks.

Data mining techniques, particularly classification algorithms, have been widely used to analyse complex data patterns and transform them into simplified, interpretable attributes [5]. By leveraging historical procurement data, classification models can identify underlying patterns and relationships

Journal of Information Systems Engineering and Management
2025, 10(52s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

among factors influencing supply timeliness [6].This study aims to bridge the existing knowledge gap by investigating the potential of data mining classification techniques in enhancing supply delay prediction for procurement processes, particularly in state-owned enterprises (SOEs) operating under strict regulatory frameworks. The study will develop and evaluate predictive models utilizing Logistic Regression, Support Vector Machine, Random Forest, and Deep Learning techniques to identify the most accurate and reliable approach for predicting supply delays.

The findings of this research are expected to contribute not only to the theoretical understanding of supply delay prediction but also to practical advancements in supply chain management. A robust predictive model will enable companies to anticipate supply delays in advance, allowing them to take proactive measures to mitigate operational disruptions and financial risks. Ultimately, this study aims to provide an innovative and data-driven approach to optimizing procurement efficiency and enhancing supply chain resilience in highly regulated industries.

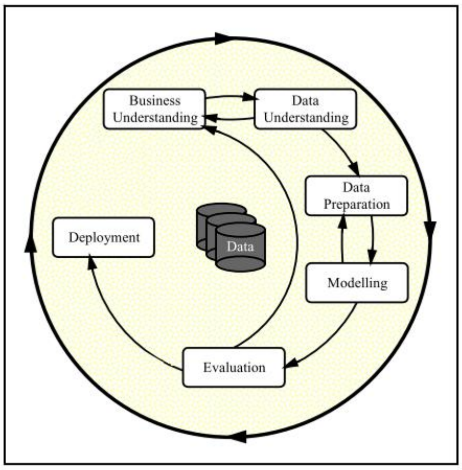
We found previous studies that predicted disruptions and risks in supply chain management (Upstream SCM) with varied and complex data patterns and there were also predictions using simple regression methods [6]. In linear regression, the relationship between independent variables and dependent variables is assumed to be linear, so the model produces predictions in the form of continuous values. Linear regression is ideal for predicting dependent variables such as income, height, or temperature, where the response is in the form of a continuous number [7]. However, when dependent variables are categorical, such as "yes/no," "true/false," or "sick/healthy," linear regression becomes unsuitable because predictions can produce values outside the range of 0 and 1, which is difficult to interpret for categories.

In this study, we made a prediction on one of the State-Owned Enterprises (SOEs) with a data mining classification technique. By simplifying the data by mapping the data variants into groups that are easier to understand, we hope that this study can contribute to the science of whether the classification method remains accurate to predict supply delays at the stage of the procurement process.

II. RESEARCH METHODS

This study adopts the CRISP-DM Method in classifying mining data. CRISP-DM (Cross-Industry Standard Process for Data Mining) is a methodology consisting of six main stages: Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, and Deployment. The phase CRISP-DM can be seen in figure 1.

FIGURE 1. CRISP-DM METHODE [8]



Procurement of goods is a systematic process that starts from the identification of needs, evaluation of budget and availability, selection of partners based on price, quality, and reputation,

Journal of Information Systems Engineering and Management
2025, 10(52s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

determination of procurement methods (tender, direct appointment, e-procurement), preparation of detailed purchase order documents, receipt and verification of goods, to payment of bills to suppliers; Meanwhile, data understanding in data mining aims to dig into the characteristics of data in depth, including its type, source, quality, format, schema, and statistical value, both of which are important for ensuring operational efficiency and accurate data analysis.

In this research prediction model, we include data from three companies' business processes. First, the purchase request process. Second, the process of issuing purchase orders and the third is the process of arrival of goods. Since both processes are executed and documented in the company's ERP system, it is also our preferred data source. The overall exported data from the company's business process transactions can be seen in table 1.

TABLE 1. OVERVIEW OF PROCUREMENT DATA

Data	Format
PR Number	Integer
Inventory Category	Text
Material Group	String
Material Number	String
Departement (Requisitioner)	String
Description	String
Quantity PR	Integer
Satuan PR	String
Estimasi PR	Integer
Currency PR	String
PO Number	Integer
Item PO	Integer
Date Ordered	Date
Del Date PO	Date
Vendor Name	String
Qty PO	Integer
Satuan PO	String
Delivery Completed	Date
Purchasing Group	String
Total Amount	Integer
Methode	String
Currency PO	String
Total Amount in Local Curr	Integer
PO Status	Text
Vendor Code	Integer
City	Text
Vendor Account Group	Integer
Incoterm	Text

Journal of Information Systems Engineering and Management
2025, 10(52s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

The phase aim to reduce errors, improve reliability, and maximize the information that can be generated from the data. The predicted variable or variable that is the target class is the status of the arrival of goods, which is obtained from the difference between the commitment date data (Delivery Date PO) and the realization data of the arrival of goods (Delivery Completed).

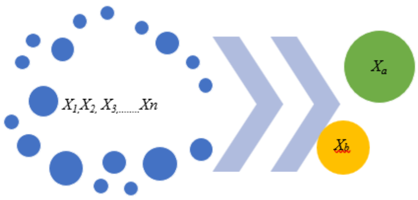


FIGURE 2. DATA REDUCTION

From the results of the reduction, the criteria for goods that fall into the category of arriving late and arriving on time are obtained. An example of an illustration can be seen in figure 2 and result data statistic of Variable prediction and target can be seen table 2.

TABLE 2. DATA STATISTICS VARIABLE

Classification Variables	**Value**	**Total**	**%**
Arrival Status	On time	8.139	72,88%
	Late	3.029	27,12%
Types of Goods	Spare parts	6.316	56,55%
	Non Sparepart	4.852	43,45%
Goods Users	Factory	6.702	60,01%
	Non Factory	4.466	39,99%
Arrival Incoterms	DDP	10.406	93,18%
	Non-DDP	762	6,82%
Procurement Methods	Tender	10.679	95,62%
	Direct (Appointment/Election)	489	4,38%

Journal of Information Systems Engineering and Management
2025, 10(52s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Vendor Conditions	Has Outstanding	9.479	84,88%
	Complete Supply	1.689	15,12%
Value Category	Low	9.162	82,04%
	High	2.006	17,96%
Inventory Category	Stock	6.702	60,01%
	Non Stock	4.466	39,99%

The process of reducing or simplifying variables to reduce the complexity of the dataset. This step aims to handle the large variety of data, both in the form of features (predictor variables) and target classes, so that the dataset becomes simpler and easier to use in classification modelling [9]. This simplification not only speeds up the analysis and modeling process but also helps to avoid problems such as overfitting and difficult interpretation for buyers in the Company.

This study is limited to the classification model. The classification model in data mining is an algorithm used to predict new data labels based on the characteristics of that data.

Logistic Regression

The logit function is a key component in logistic regression that plays a role in converting the output of a linear model into a probability that ranges between 0 and 1. In ordinary linear regression, linear combinations of independent variables are generated in an infinite range of values, both negative and positive. However, since logistic regression aims to predict binary outcomes e.g., in this study is data on goods arriving "late" or "on time", we need a way to map these outputs to a range of probabilities. This is where the logit function comes into play.

TABLE 3. COMPARISON LINEAR & LOGISTIC REGRESSION

Aspects	**Linear Regression**	**Logistic Regression**
Curve Shape	Straight (linear)	Sigmoid Curve (S-shaped)
Equation	$y = mx + c$	$P(y) = \frac{1}{1 + e^{-(mx+c)}}$
Output	Continuous value (can be negative or positive)	Probability (between 0 and 1)
Interpretation of Output	Directly used as a prediction result	Converted to a label (0 or 1) using a threshold (e.g., 0.5)

Journal of Information Systems Engineering and Management
2025, 10(52s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Application	Predicting the value of prices, temperatures, revenues, etc.	Predicting class categories
Model Type	Regression (continuous value prediction)	Classification (category prediction)

Support Vector Machine (SVM)

SVM effectively separates data into two distinct classes based on position relative to the optimal hyperplane. In this study, we will also use SVM for problems with non-linear data. The modeling process of the Support Vector Machine can be seen in figure 3 as follows:

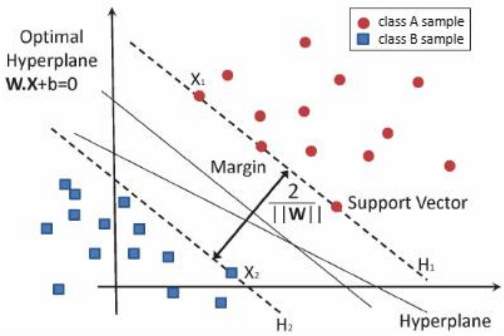


FIGURE 3. SVM MODELING [10].

Random Forest

Random Forest is built using bootstrap sampling techniques, which select a random subset of the training data for each decision tree. In addition, each tree uses only a different subset of features on each branch, thus reducing the risk of overfitting and making the model more generalized. When receiving new data, each Decision Tree in Random Forest will provide a prediction of whether the delivery will be on time or delayed. The final prediction is determined by the majority voting method, where if the majority of the trees predict "late", then the model will classify the order as at risk of delay. Conversely, if the majority of trees predict "on time", then delivery is most likely not a problem. The modeling process of the Random Forest can be seen in figure 4 as follows:

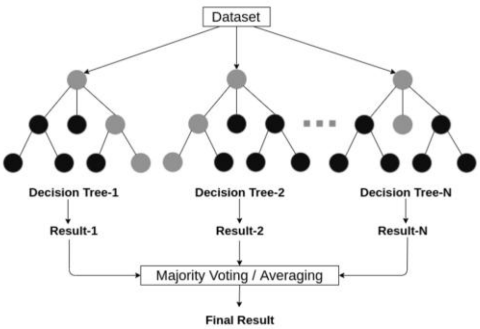


FIGURE 4. RANDOM FOREST STRUCTURE [11].

Deep Learning

One of the deep learning models that can be applied in this case is Feedforward Neural Networks (FNNs), which are basic artificial neural networks in which data flows unidirectionally from the input layer to the hidden layer all the way to the output layer. In the early stages, FNN requires historical data on supply delays. This data is used as an input variable, while the target variable is a delay category, e.g.

Journal of Information Systems Engineering and Management
2025, 10(52s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

"on time" or "late". After the data is prepared, the model is trained with backpropagation and gradient descent methods to adjust the weights and biases in each neuron, so that the model can recognize patterns in the data. The modeling process of the Deep Learning can be seen in figure 5 as follows:

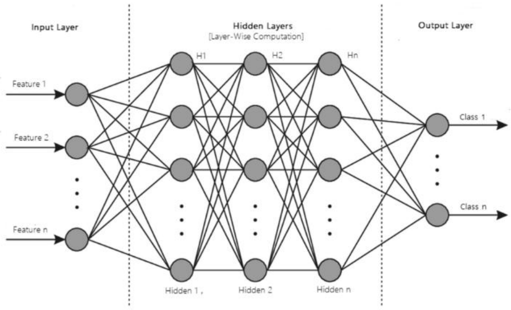


FIGURE 5. DEEP LEARNING STRUCTURE [11]

The right choice of metric depends largely on the context of the problem and the purpose of the analysis. If the cost of false positives and false negatives is both high, then F1-score is a good choice. However, if one type of error is more impactful, then precision or recall can be a more relevant metric. In addition, in the case of unbalanced classes, metrics such as precision, recall, and F1-score will provide more comprehensive information compared to accuracy. Look table 4

TABLE 4. MODELING PERFORMANCE PARAMETERS

Metric	**Formula**	**Explanation**
Accuracy	$= \frac{TP + TN}{TP + TN + FP +}$	The proportion of correct predictions against the entire data.
Error	$= \frac{FP + FN}{TP + TN + FP +}$	The proportion of incorrect predictions against the entire data. The smaller the error value, the better the model will perform.
Precision	$= \frac{TP}{TP + FP}$	The proportion of correct positive predictions to all positive predictions.
Recall	$= \frac{TP}{TP + FN}$	The proportion of correct positive predictions against all actual positive data.
F1-Score	$= 2\,X\left(\frac{Precision}{Precision}\right.$	Harmonic average of precision and recall. Useful when precision and

		recall are equally important.
AUC	Area under the ROC curve formed by the True Positive Rate (TPR) and False Positive Rate (FPR)	A metric that measures the model's ability to distinguish between positive and negative classes. AUC = 1 indicates a perfect model, while AUC = 0.5 indicates a random model.

In the performance assessment of the model that has been built at the modeling stage. This evaluation is important to ensure the accuracy of the model, i.e. how well the model can predict, classify, or group data and compare the performance of different models to choose the best model.

Deployment

After the machine learning model has been trained and evaluated, the next stage is to simulate predictions. This simulation aims to see how our model works on new data that has never been seen before so that its application can help companies in decision-making.

RESULTS AND DISCUSSION

The performance evaluation results of the four classification models showed diverse performance. The Deep Learning and Random Forest models emerged as the best-performing models, both recording an accuracy of 92.3%, an error rate of 7.7%, but the Deep Learning modeling had a slight advantage in the AUC value parameter of 97.7% while the Random Forest was 97.4%. The precision, recall, and F-Score measurements in both models were also excellent, demonstrating the model's ability to correctly classify samples, especially in identifying positive categories. Look table 5.

TABLE 5. MODELING PERFORMANCE PARAMETERS

Model	Acc.	Err.	Preci.	AUC	Recall	F1
RL	90.7 %	9,3%	89%	96,6 %	99,6 %	94 %
SVM	90.3 %	9,7%	90,3 %	96,6 %	97,3 %	93 %
RF	92.3 %	7,7%	90,5 %	97,4 %	100 %	95 %
DL	92.3 %	7,7%	90,5 %	97,7 %	100 %	95 %

The Logistic Regression model shows quite good performance with an accuracy of 90.7% and an AUC value of 96.6%. Nonetheless, its performance is slightly below Deep Learning and Random Forest. The SVM model has a relatively lower performance than other models, especially in the recall value, which indicates the potential for classification errors in the positive category. Overall, the comparison of ROC (Receiver Operating Characteristic) in each modeling can be depicted on the curve. Look Figure 6.

Journal of Information Systems Engineering and Management
2025, 10(52s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

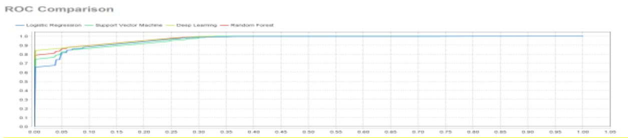


FIGURE 6. ROC CURVE

The correlation analysis of various variables across four different models revealed diverse outcomes. The variables "Price Category" and "Incoterms Arrival" consistently exhibited a significant positive correlation in most models, particularly in Logistic Regression, with correlation values of 0.602 and 0.477, respectively. Look Table 6.

TABLE 6. MODELING PERFORMANCE PARAMETERS

Variabel	RL	SVM	RF	DL
Price Category	0,602	0,078	0,446	0,423
Arrival of Incoterms	0,477	0,100	0,567	0,546
Vendor	0,088	0,049	0,075	0,101
Inventory Category	0,157	0,014	0,018	0,008
Types of Goods	0,039	0,171	0,051	0,027
Users	0,110	0,012	0,208	0,193
Procurement Methods	0,013	0,013	0,005	0,008

This indicates that these two variables have a strong influence on predictive outcomes. Other variables, such as "Vendor" and "User", also showed positive correlations, though their strengths varied across models. Deep Learning and Random Forest models tended to exhibit more diverse correlations with different variables, whereas the Support Vector Machine (SVM) model generally demonstrated lower correlations across all predictor variables. Furthermore, as presented in Table 4.9, "Incoterms Category" emerged as the most influential variable in the Random Forest and Deep Learning models, highlighting the need for improvements in Incoterms determination for each purchase order. Additionally, "Purchase Order Value Category" played a crucial role in these models, where lower-value purchase orders (below IDR 50,000,000) exhibited a higher probability of on-time arrival. Although the subject of this study involved several state-owned enterprises (SOEs) whose procurement methods are regulated by government policies, the findings indicate that the procurement method variable had the least impact compared to other variables.

CONCLUSION

This study concludes that Random Forest and Deep Learning models demonstrate the highest accuracy, lowest error rates, perfect recall, and an optimal balance between precision and recall, making them the most effective models for supply delay prediction. Deep Learning achieves the highest AUC (97.7%), indicating superior performance in distinguishing between on-time and delayed deliveries, particularly in handling imbalanced datasets. The findings highlight that Incoterms and Purchase Order Value are the most influential factors in supply delay prediction, whereas the procurement method has minimal impact. The study also reinforces the effectiveness of the CRISP-DM framework, demonstrating its capability to improve predictive accuracy in complex business processes. From a practical perspective, integrating these predictive models into supply chain management systems can enable early detection of potential delays, enhancing procurement efficiency. Future research should explore model applicability in non-state-owned enterprises, integrate predictive models with ERP systems, conduct continuous performance monitoring, and utilize larger, more diverse datasets to enhance model generalization by incorporating external factors such as economic conditions, natural disasters, and policy changes.

AUTHOR CONTRIBUTION

Writing—original draft, Kasthalani; Methodology, Kasthalani.; Data curation, Kasthalani; Review and editing, M. Z.; Providing resources, M. Z. All authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] Sunil. Chopra and Peter. Meindl, *Supply chain management : strategy, planning, and operation*. Pearson, 2016.

[2] D. Soelistijanto and H. Siringoringo, "Developing information system/technology strategy in supporting business process: a case study of insurance state-owned enterprises," *Int J Bus Inf Syst*, vol. 42, no. 1, pp. 43–71, 2023, doi: 10.1504/IJBIS.2020.10030464.

[3] B. Ratiko and H. Setiyawati, "The Effect Of Transparency Principles And E-Procurement On The Function Of The Internal Control Systems And The Impact On Financial Accountability," *International Journal of Management Studies and Social Science Research*, vol. 4, no. 2, pp. 161–175, 2022, [Online]. Available: www.ijmsssr.org

[4] Perpres Nomor 12, "Peraturan Presiden Republik Indonesia Nomor 12 Tahun 2021 Tentang Perubahan Atas Peraturan Presiden Nomor 16 Tahun 2018 Tentang Pengadaan Barang/Jasa Pemerintah," *Republik Indonesia*, no. 086130, pp. 1–47, 2021.

[5] G. S. Sajja, H. Pallathadka, K. Phasinam, and S. Ray, "Using Classification Data Mining for Predicting Student Performance," *ECS Trans*, vol. 107, no. 1, p. 10217, 2022, doi: 10.1149/10701.10217ecst.

[6] F. Steinberg, P. Burggräf, J. Wagner, B. Heinbach, T. Saßmannshausen, and A. Brintrup, "A novel machine learning model for predicting late supplier deliveries of low-volume-high-variety products with application in a German machinery industry," *Supply Chain Analytics*, vol. 1, no. February, p. 100003, 2023, doi: 10.1016/j.sca.2023.100003.

[7] Hosmer, *Applied Logistic Regression.3rd edn John New York: Wiley;* 2013.

[8] R. Wirth and J. Hipp, "CRISP-DM: towards a standard process model for data mining. Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining, 29-39," *Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining*, no. 24959, pp. 29–39, 2000, [Online]. Available: https://www.researchgate.net/publication/239585378_CRISP-DM_Towards_a_standard_process_model_for_data_mining

[9] W. Jia, M. Sun, J. Lian, and S. Hou, "Feature dimensionality reduction: a review," *Complex and Intelligent Systems*, vol. 8, no. 3, pp. 2663–2693, Jun. 2022, doi: 10.1007/s40747-021-00637-x.

[10] E. García-Gonzalo, Z. Fernández-Muñiz, P. J. G. Nieto, A. B. Sánchez, and M. M. Fernández, "Hard-rock stability analysis for span design in entry-type excavations with learning classifiers," *Materials*, vol. 9, no. 7, pp. 1–19, 2016, doi: 10.3390/ma9070531.

[11] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Comput Sci*, vol. 2, no. 3, pp. 1–21, 2021, doi: 10.1007/s42979-021-00592-x.