Jurnal Informatika Ekonomi Bisnis

http://www.infeb.org

2024 Vol. 6 Iss. 4 Hal: 732-739 e-ISSN: 2714-8491

Comparative Analysis of Multicriteria Inventory Classification and Forecasting: A Case Study in PT XYZ

Dhanang Jaka Purwandaru[1]✉, Yova Ruldeviyani[2], Sani Novi Nugraheni[3], Galuh Prisillia[4]

[1,2,3,4]Universitas Indonesia

dhanang.jaka@ui.ac.id

Abstract

One crucial aspect of supply chain management is inventory management. Inefficient inventory management can lead to various issues, such as product expiration, where a high number of items in the warehouse either have expired or are approaching expiration. This issue is experienced by a distribution SME in Indonesia, PT XYZ. Without such classifications, it becomes challenging to predict demand and manage stock levels efficiently. Therefore, the aim of this study is to classify inventory to identify the most important items to business and make a forecasting model of sales quantity to predict inventory replenishment using machine learning algorithms. To advance our research, we adopted the Cross Industry Standard Process for Data Mining (CRISP-DM) methodology. For inventory classification, we conducted a hybrid approach that combined TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) and ABC analysis (A: high-value items, B: medium-value items, and C: low-value items). The data employed in this study comprised secondary data, including purchase orders, sales orders, and stock movement records. The result reveals that 11 of the total 383 items under class A are important items for business. After obtaining labels from the ABC Analysis, we proceed to train models using KNN, SVC, and Random Forest for predicting inventory classification. Notably, the Random Forest model showcased remarkable performance and outperformed the rest of the models, achieving an accuracy of 99.21%. For inventory forecasting ARIMA displays a competitive performance with RMSE value 5.305 and MAE value 3.476, indicating a relatively accurate prediction with lower forecasting errors than two other models.

Keywords: Inventory Management, Small and Medium-sized Enterprise, Technique for Order Preference by Similarity to Ideal Solution, ARIMA, LSTM.

1. Introduction

Effective organizational operations hinge on the critical role of supply chain management, ensuring profitable processes. It involves emphasizing market supply and demand dynamics to coordinate seamlessly with suppliers and acquire high-quality materials [1]. To emphasize market supply and demand dynamics, we need inventory control. Inventory control is the heart of operations management, where replenishment actions must be taken to minimize costs [2]. The goal of inventory control is to assess and manage an inventory level that reduces overall system and organizational costs or costs associated with industrial factories [3]. Nevertheless, there are several problems with inventory control. In most real-world inventory control problems, demand changes over time and the true underlying demand distribution are never fully known to the inventory manager. The manager makes dual use of historical demand data to populate the current demand distribution and to detect fundamental changes in the demand-generating process [4].

Inventory is one of the most important assets that companies have. The turnover of inventory represents one of the primary sources of revenue generation and subsequent earnings for the company, namely inventory turnover ratio. Inventory refers to the goods or materials used by a company for production or sale purposes [5]. The nature of inventory is different from one organizational sector to another, depending on the industry, for instance, a manufacturer's inventory is different from that of a retailer or service provider [6].

Due to its importance for business continuity, good inventory management is essential to maintaining inventory at the most favorable costs. Company must strategically plan, organize, and meticulously control inventory processes [7], including controlling turnover rate and managing replenishment. The effective inventory management results the reduction of storage and holding costs, while fully leveraging all available supply capabilities. Thus, it can lead to improved cash flow, increased sales, and the prevention of loss from overstocked or understocked items.

A case study related to inventory management optimizing will be conducted in PT XYZ which is a Small and Medium-sized Enterprise (SME) company, operating in the distribution sector in Indonesia. This company has implemented enterprise resource planning (ERP) in its business processes since February 2023. After implementation, the ERP still requires further development since there are still some constraints. Firstly, there is a concerning issue with product expiration, as a high number of products in the

Accepted: 24-10-2024 | Revision: 01-11-2024 | Publication: 31-12-2024 | doi: 10.37034/infeb.v6i4.1014

warehouse have either expired or are nearing expiration. This indicates inefficiencies in inventory turnover and management. Secondly, the stock availability for high-demand products is a persistent problem, leading to potential lost in sales and customer dissatisfaction. Furthermore, the absence of product classification into categories hampers effective inventory control and management. Without such classifications, it becomes challenging to predict demand and manage stock levels efficiently.

Those hurdles could significantly impact the company's operations and profitability if not adequately addressed. Ineffective inventory management can lead to various challenges, such as stockouts of high-demand products, overstock of less popular items, and financial losses due to expired or obsolete stock. These issues not only affect the company's bottom line but also its reputation and customer satisfaction levels. Moreover, without proper inventory control, the company may struggle to respond effectively to market demands, leading to missed opportunities and decreased competitive advantage. Therefore, addressing these inventory constraints is crucial for maintaining operational efficiency, ensuring customer satisfaction, and sustaining the overall health of the business.

The digital revolution and machine learning contribute to organizational operations by establishing an efficient network connecting organizations and suppliers, facilitating the coordination of market demands and supplies. The potential of digitization in inventory management lies in establishing a framework built on precise market predictions [8]. Therefore, to address challenges in inventory management in PT XYZ, the aim of this study is to classify inventory to identify the most important items to business and make a forecasting model of sales quantity to predict inventory replenishment using machine learning algorithms, thus there are two research questions need to be addressed, as follows:

RQ1: How can inventory be classified to identify the most important items for a business?

RQ2: How can sales quantity of an item be forecasted to predict inventory replenishment?

Prior research in inventory classification has explored diverse methodologies, as demonstrated by Roy et al., who applied the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) in multicriteria decision making (MCDM) followed by ABC analysis. Their investigation incorporated K-nearest neighbors (KNN) and support vector machine (SVM) implementations, with results indicating superior test accuracy for the KNN model over the SVM model. Similarly, Khanokar and Kane utilized the A. Hadi-Vencheh model in MCDM for inventory classification, complemented by the K-Means clustering algorithm. However, these studies did not incorporate sales forecasting to predict organizational inventory requirements. In contrast, our research focused on inventory classification using Support Vector Classifier (SVC), KNN, and Random Forest, with an additional emphasis on sales forecasting using ARIMA, SARIMA, and LSTM models to enhance organizational inventory management.

2. Research Method

In the Figure 1, we present a visualization of our research methodology, which combines the CRISP-DM methodology with various data modeling and evaluation techniques tailored to inventory classification and forecasting.

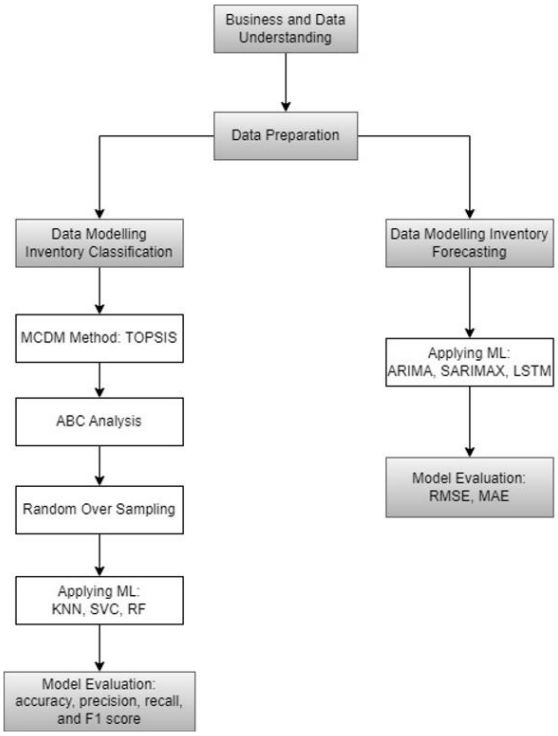

Figure 1. Research Methodology

CRISP-DM methodology is a widely recognized standard process model that outlines common techniques and phases in data mining [9]. This model offers a methodical framework for data mining projects, ensuring a systematic and efficient progression from understanding the business problem to the deployment of solutions based on data analysis.

CRISP-DM consists of six phases, starting with business understanding, which focuses on understanding project objectives and requirements from a business perspective [9]. The data understanding phase initiates with initial data collection, progressing through activities to familiarize oneself with the data, identify data quality issues, gain initial insights, and detect subsets for forming hypotheses about hidden information [10]. The data preparation phase encompasses all activities to construct the final dataset that will be input into

modeling tools from the acquired data [11]. In the modeling phase, various modeling techniques are selected and applied, with their parameters calibrated to optimal values. The evaluation phase involves a more thorough assessment of the model, reviewing the steps taken to construct it to ensure it effectively meets the business objectives. Finally, the deployment stage entails integrating the chosen model into the business environment for informed decision-making.

2.1 Business and Data Understanding

In our research, we conduct both business and data understanding to ensure alignment with broader business objectives, as emphasized in the business understanding phase of CRISP-DM [12]. We enhance our comprehension of the organization's operations by engaging in interviews with its employees, which aids in obtaining a comprehensive understanding of the business processes, particularly in inventory management. One of the most important aspects of inventory management is inventory replenishment. This is the operational process of ordering and restocking goods and materials, aiming to keep inventory levels within a range that satisfies customer demand without excess [7]. Understanding this process is crucial for aligning our data analysis with the organization's needs.

Subsequently, we gather the requisite data for our research and strive to gain deeper insights into it. We obtained a dataset comprising 200,000 rows of historical inventory data, including product ID, date, product category, purchase quantity, sales quantity, and selling price. Throughout this process, we conducted meticulous exploratory data analysis to examine the significance of each feature within the dataset. Furthermore, the insights gained from employee interviews and exploratory data analysis contribute to refining our understanding of both the business context and the intricacies of the dataset. This holistic understanding lays a robust foundation for subsequent stages in our research, facilitating more informed decision-making throughout the analysis and modeling processes.

2.2 Data Preparation

The initial dataset consists of transaction data from PT XYZ, which we preprocess using various methods. After collecting the data, we proceed to the next stage involving cleaning and normalization. During this phase, we refine the data to eliminate any irrelevant or unnecessary information, thereby enhancing the focus of the research. Categorical data is transformed into numerical values to ensure compatibility with our machine learning model. We also pay meticulous attention to addressing missing data values, and data standardization is performed to enhance compatibility [13]. Concurrently, a comprehensive analysis is conducted to identify and manage outliers, safeguarding the integrity of the dataset.

The processed dataset is then grouped based on date and product ID. Subsequently, we sum the inventory, resulting in new features such as annual buy quantity, annual sell quantity, annual product price, annual unit cost, and lead time for each product. These features are validated with subject matter experts from PT XYZ to obtain weighting for classification using TOPSIS. Afterwards, the prepared data is partitioned into distinct sets for training and testing.

2.3 Modelling

The primary focus of this study is to evaluate the significance of inventory by utilizing TOPSIS for categorization, followed by ABC classification. TOPSIS introduced by Hwang and Yoon in 1981, is a fundamental method that useful in situations where decision-makers need to choose the best option from a set of multiple or conflicting criteria [14]. TOPSIS requires establishing specific weight values for attributes prior to performing the calculations [15]. TOPSIS then identifies the ideal and anti-ideal solutions. The next step is calculating the Euclidean distance of each alternative from these ideal and anti-ideal points [16]. The final ranking of alternatives is based on their relative closeness to the ideal solutions.

Subsequently, we employ machine learning algorithms SVC, KNN, and random forest to develop a model for improved inventory categorization. SVC, a variant of Support Vector Machine, excels at classifying high-dimensional data by transforming input data into a higher-dimensional space, effectively managing nonlinear relationships [17]. KNN is an instance-based learning algorithm that predicts the classification of a test sample based on the majority class among its *k* nearest neighbors [18]. Random Forest is an ensemble learning method that combines multiple decision trees to enhance prediction accuracy and prevent overfitting by averaging the results from trees built on random subsets of data [19].

To accomplish this goal, we selected a dataset containing 383 inventory items, focusing on four key attributes: annual sales, inventory price, annual unit cost, and lead time. These attributes are crucial in the inventory item classification process and form the basis for the multi-criteria decision-making (MCDM) approach using TOPSIS to guide the ABC analysis. This process involves organizing the data using TOPSIS and classifying the inventory items through the ABC classification method.

Another significant aspect of the study involves forecasting inventory replenishment for selected products. In this endeavor, we leverage time series forecasting models such as ARIMA, SARIMA, and LSTM networks to make predictions based on the most recent data. ARIMA models are well-regarded for their ability to model various data series by relying on their own past values, trends, seasonality components, and

error terms [20]. They can be extended to include exogenous variables and seasonal components, evolving into SARIMAX models. LSTM networks, a type of recurrent neural network, have revolutionized sequence prediction due to their capacity to remember long-term dependencies [21].

The application of forecasting methods on time series data also involves exponential smoothing models, which generate projections as the weighted average of previous data. Various approaches to exponential smoothing include simple, double, and triple exponential smoothing models. The ARIMA model combines the best aspects of autoregressive and moving average models by addressing the differentiation of time series data. It iteratively executes three stages using the Box-Jenkins method: model identification, parameter estimation, and diagnostic checking.

2.4 Evaluation

For assessing model performance, we employed metric such as accuracy, precision, recall and F1 score and confusion matrix for classification These metrics provide a comprehensive evaluation of the model's predictive capabilities. The equations for these indicators can be seen on Equation 1 until Equation 4.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (4)$$

The Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) will be used for forecasting evaluation. The root mean square error (RMSE) has been used as a standard statistical metric to measure model performance [22]. MAE is widely used in machine learning models to measure its performance evaluating the difference between the predicted values and the real ones if they are at the same scale [23]. A lower score on these metrics signifies enhanced model performance, which can be considered for implementation in a company. The equation for above indicators can be seen on Equation 5 and Equation 6.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (5)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (6)$$

The deployment phase represents a stage wherein the data mining results are delivered into the operational environment [24]. However, in our research, the deployment phase was not undertaken. This decision was made on the grounds that the scope of our study did not encompass the deployment stage. The primary focus of our research was confined to the stages of understanding the problem, data preparation, and the modeling process, with particular emphasis on inventory classification and forecasting using a set of mathematical algorithms. Therefore, the study was concluded post the evaluation stage, without progressing to the deployment phase

3. Result and Discussion

3.1 Inventory Classification

After studying the characteristics of the business, reviewing historical data obtained from the organization, and conducting interviews with the employees, the authors initiate the data preparation, modeling, and evaluation phases. As mentioned earlier, we obtained approximately 200,000 rows of historical data. We conducted feature engineering on the original data, resulting in 383 unique product IDs with columns for Annual Sales, Product Price, Annual Unit Cost, and Lead Time as shown in Table 1. These specific features are analyzed using the TOPSIS method.

Table 1. Data characteristics

Input Feature	Description	Scale
Annual Sales	The sum of annual sales of each product	numerical value > 0
Product Price	The price of each product	numerical value > 0
Average Unit Cost	The average of utilities cost that divided into each product	numerical value > 0
Lead Time	Average length of time between when the item order and when it is received	numerical value > 0

Before conducting an ABC Analysis on the inventory, we performed inventory ranking using TOPSIS scores. The initial step in this stage is to determine the weights for the analysis, which were obtained through interviews with the organization's management. The weights utilized in this study are as follows: Annual Sales (0.45), Product Price (0.30), Average Unit Cost (0.125), and Lead Time (0.125).

The initial step in the TOPSIS analysis involves normalizing the decision matrix [25]. In this matrix, each row corresponds to an alternative (such as an inventory item), and each column corresponds to a criterion (e.g., Annual Sales, Product Price, Average Unit Cost, Lead Time). The values in the matrix indicate the performance of each alternative concerning each criterion. The normalization process is crucial as it scales these values, ensuring that all criteria are on a comparable scale. This step is vital because different criteria may have diverse units or ranges. The normalize decision matrix then assigned weight based on its relative importance based on professional judgment.

The result of these steps is the weighted normalized decision matrix.

The computation of separation measures in TOPSIS involves determining the distance between each alternative and both the ideal solution and the anti-ideal solution. The ideal solution serves as a benchmark, representing the optimal values for each criterion Conversely, the anti-ideal solution represents the worst values for each criterion. Commonly, the Euclidean distance is employed to calculate these separation measures. For each alternative, the distance to the ideal solution and the distance to the anti-ideal solution are computed.

For instance, the provided values for the ideal solution are 0.2209, 0.0641, 0.0644, 0.0081, while the values for the anti-ideal solution are 0.0, 0.0, 1.153e-06, 0.0041. These values set the reference points for assessing how well each alternative aligns with the optimal and suboptimal criteria. The calculated distances, derived through the Euclidean distance formula, will help quantify the relative proximity of each alternative to the ideal and anti-ideal solutions in the TOPSIS analysis.

After that the steps is to calculate the relative closeness of the ideal and anti-ideal solution with each inventory. Calculated by dividing the distance to the anti-ideal solution by the sum of the distances to both the ideal and anti-ideal solutions, this metric offers a comprehensive evaluation of how well each alternative aligns with the optimal and suboptimal criteria. In the final step of the analysis, the alternatives, in this case, inventory items, are ranked based on their relative closeness values. The alternative with the highest relative closeness is deemed the most favorable or preferred. This systematic approach aids in prioritizing and ranking inventory items, considering the significance of each criterion and the distances to both ideal and anti-ideal solutions.

After applying Pareto's principle, the existing inventory underwent a systematic reorganization, arranged in a descending order. Following this, the items were assigned to A, B, and C inventory classes using the results obtained through the TOPSIS method.

The ABC analysis output, detailed in Figure 2 and Figure 3, reveals the classification of 11 items under class A, 33 items under class B, and 339 items under class C.

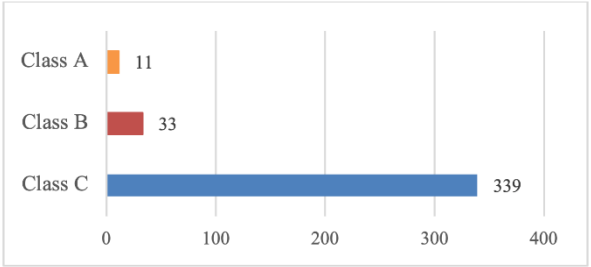

Figure 2. ABC Analysis Output

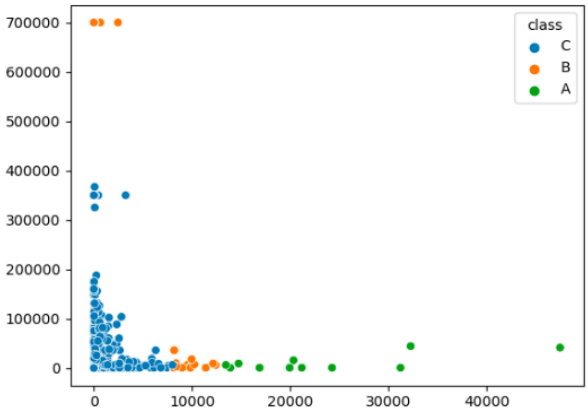

Figure 3. Plot of Classification Based on ABC Analysis

After obtaining labels from the ABC Analysis, we proceed to train models using KNN, SVC and Random Forest for improved inventory classification. In this research, KNN and SVC and Random Forest are employed to build the models. Prior to model training, we employ random oversampling to generate additional data for both the training and testing stages. Specifically, random oversampling is utilized, with 200 items assigned to class A, 300 items to class B, and 500 items to class C.

We can view comprehensive overview of the performance metrics for three distinct machine learning models—SVC, KNN, and random forest applied to the analyzed dataset in Table 2.

Table 2. Model Result

Model	Accuracy	Precision	Recall	F1 Score
SVC	0.9556	0.9578	0.9554	0.9465
KNN	0.9843	0.9818	0.9843	0.9816
RF	0.9992	0.9974	0.9921	0.9944

The models were evaluated based on key performance indicators Accuracy, Precision, Recall, and F1 Score. The SVC model demonstrated commendable results with an accuracy of 95.56%, precision of 95.78%, recall of 95.54%, and an F1 Score of 94.65%. The KNN model exhibited even higher performance, achieving accuracy of 98.43%, precision of 98.18%, recall of 98.43%, and an F1 Score of 98.16%. Notably, the random forest model showcased remarkable performance and outperformed the previous models, achieving accuracy of 99.21%, precision of 99.74%, recall of 99.21%, and an F1 Score of 99.44%. These metrics collectively underscore the effectiveness of the models in learning from the training data and generalizing well to new, unseen data, highlighting their robustness in classification tasks. From the confusion matrix in Figure 4, it is evident that the random forest model exhibits fewer false predictions compared to other models. Figure 4 illustrates that the random forest model misclassifies instances of class A as class B with an error rate of 8.33%. This error rate is notably lower than the inaccuracies observed in the predictions made by the

KNN and SVC models as illustrated in Figure 5 and Figure 6.

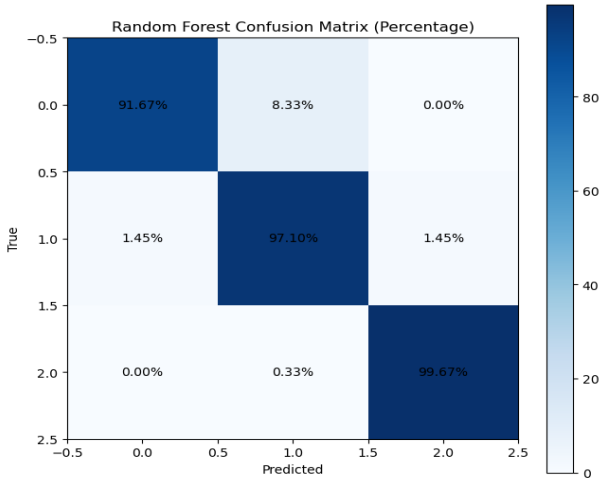

Figure 4. Confusion Matrix Random Forest

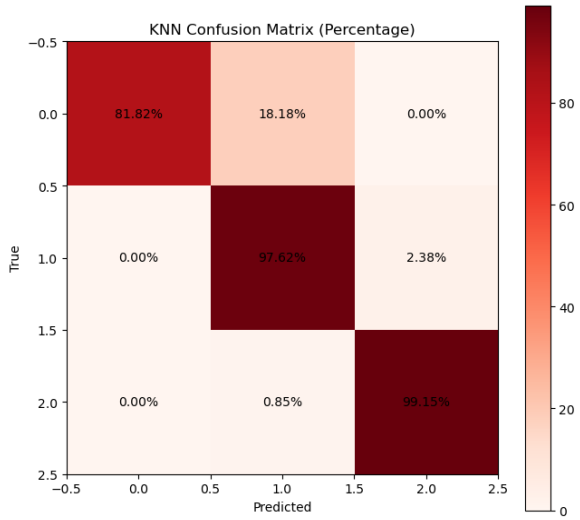

Figure 5. Confusion Matrix KNN

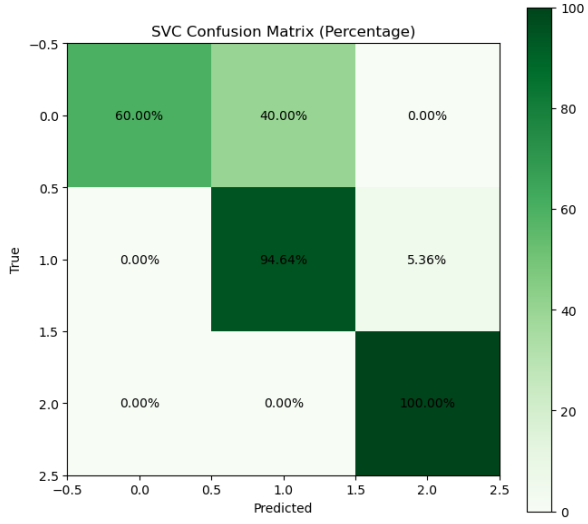

Figure 6. Confusion Matrix SVC

3.2 Inventory Forecasting

The authors sought to explore the optimal inventory level that would maximize sales and minimize inventory costs within the organization. To address these challenges, we employed time series forecasting to model the inventory stock. Specifically, a stock from class A inventory was selected as a representative case to demonstrate the efficacy of machine learning models in predicting inventory levels, thereby aiding organizations in making informed purchasing decisions. The forecasting process commenced with data preprocessing, wherein we meticulously prepared the selected inventory's data.

The authors divided the time series into two parts, allocating 80% for training the model and reserving 20% for testing its performance. Another step taken by the authors involved indexing the data using the date column of the dataset. Subsequently, the prepared data was fed into the model, and various models were evaluated to determine the most effective one for forecasting.

The historical and predicted inventory using ARIMA are illustrated in Figure 7, while Figure 8 showcases the results using SARIMAX. The outcomes using LSTM are depicted in Figure 9.

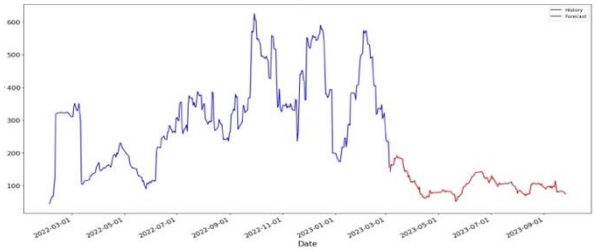

Figure 7. ARIMA Result

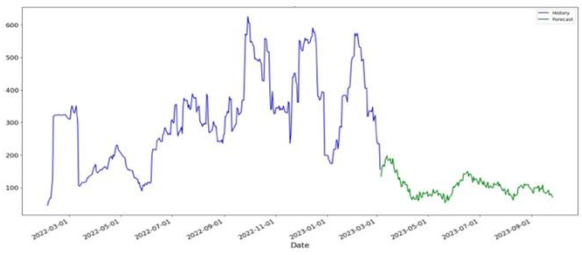

Figure 8. SARIMAX Result

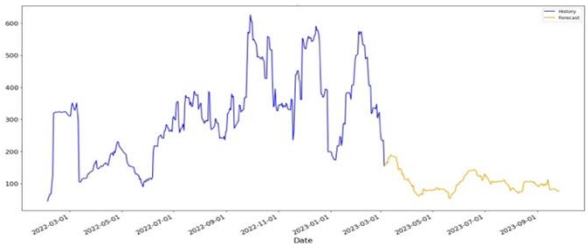

Figure 9. LSTM Result

The evaluation of three distinct time series forecasting models—ARIMA, SARIMAX, and LSTM—is

summarized in the Table 3, presenting key performance metrics for each model.

Table 3. Evaluation model

Model	RMSE	MAE
ARIMA	5.305	3.476
SARIMA	7.310	5.659
LSTM	13.113	12.296

The ARIMA model displays a competitive performance with a RMSE of 5.305 and a MAE of 3.476, indicating a relatively accurate prediction with lower forecasting errors. In contrast, the Seasonal Auto Regressive Integrated Moving Average with Exogenous Factors (SARIMAX) model demonstrates slightly higher errors, as reflected in an RMSE of 7.310 and an MAE of 5.659. Notably, the Long Short-Term Memory (LSTM) model, a recurrent neural network, exhibits the highest forecasting errors among the three models, with an RMSE of 13.113 and an MAE of 12.296. While the ARIMA model proves to be the most accurate in this comparison, the selection of the optimal model depends on specific data characteristics and the desired trade-off between accuracy and computational complexity.

This study effectively demonstrates the use of TOPSIS in multi-criteria decision making, particularly for inventory management with multiple attributes. It also explores the enhancement of traditional inventory classification methods through machine learning models such as KNN, SVC, and RF, potentially inspiring further research into their adaptability across various inventory contexts. The study's findings on ARIMA's effectiveness in forecasting inventory needs contribute to the broader discussion on suitable time series models for predicting business metrics. Organizations can use these results as a structured approach to inventory classification and forecasting, benefiting from data-driven methods like machine learning and time series forecasting to manage inventories more accurately and reliably, thus potentially reducing costs and improving supply chain efficiency. The proposed model can be applied to different industries or business scales, offering an alternative approach to inventory management, especially in classification and prediction. However, the study has limitations, as its results are based on specific models and datasets, which may vary in performance and applicability across different inventory types or business environments. The analysis is heavily dependent on data quality and comprehensiveness, with incomplete or inaccurate data significantly impacting results. A larger dataset might yield better outcomes. Additionally, the use of random oversampling to address class imbalances can introduce biases, affecting model performance.

4. Conclusion

The comprehensive analysis undertaken in this study encompasses various stages, starting from the initial exploration of historical data, feature engineering, and inventory ranking using the TOPSIS method to the subsequent application of ABC Analysis and machine learning models for improved inventory classification. The study exemplifies the application of time series forecasting on a representative inventory item from the A class inventory items. Three model were examined, ARIMA, SARIMAX, and LSTM, followed by evaluation using RMSE and MAE. The study results show that ARIMA has the best performance among the models tested with RMSE value 5.305 and MAE value 3.476, indicating a relatively accurate prediction with lower forecasting errors than two other models. The results of this study present a systematic method for inventory classification and forecasting that can be used by the organization. Incorporating data-driven techniques such as machine learning models and time series forecasting provides a more precise and dependable approach for businesses in handling their inventories. This approach has the potential to lower costs and enhance the efficiency of supply chain operations. Furthermore, future research could involve applying these methods to a wider range of datasets to test their effectiveness and adaptability across different inventory types and business models, investigate other machine learning models to enhance model performance, especially in the context of inventory classification and inventory forecasting.

References

[1] Wisetsri, W., Donthu, S., Mehbodniya, A., Vyas, S., Quiñonez-Choquecota, J., & Neware, R. (2022). An investigation on the impact of digital revolution and machine learning in supply chain management. *Materials Today: Proceedings*, *56*, 3207-3210. https://doi.org/10.1016/j.matpr.2021.09.367

[2] Boute, R. N., Gijsbrechts, J., Van Jaarsveld, W., & Vanvuchelen, N. (2022). Deep reinforcement learning for inventory control: A roadmap. *European Journal of Operational Research*, *298*(2), 401-412. https://doi.org/10.1016/j.ejor.2021.07.016

[3] Harifi, S., Khalilian, M., Mohammadzadeh, J., & Ebrahimnejad, S. (2021). Optimization in solving inventory control problem using nature inspired Emperor Penguins Colony algorithm. *Journal of Intelligent Manufacturing*, *32*, 1361-1375. https://doi.org/10.1007/s10845-020-01616-8

[4] Wang, Z., & Mersereau, A. J. (2017). Bayesian inventory management with potential change-points in demand. *Production and Operations Management*, *26*(2), 341-359. https://doi.org/10.1111/poms.12650

[5] Islam, S. S., Pulungan, A. H., & Rochim, A. (2019, December). Inventory management efficiency analysis: A case study of an SME company. In *Journal of Physics: Conference Series* (Vol. 1402, No. 2, p. 022040). IOP Publishing. https://doi.org/10.1088/1742-6596/1402/2/022040

[6] Munyaka, J. B., & Yadavalli, V. S. S. (2022). Inventory management concepts and implementations: a systematic review. *South African Journal of Industrial Engineering*, *33*(2), 15-36. https://doi.org/10.7166/33-2-2527

[7] Teplická, K., & Čulková, K. (2020). Using of optimizing methods in inventory management of the company. *Acta logistica*, *7*(1), 9-16. https://doi.org/10.22306/al.v7i1.150

[8] Hartley, J. L., & Sawaya, W. J. (2019). Tortoise, not the hare: Digital transformation of supply chain business processes. *Business Horizons*, *62*(6), 707-715. https://doi.org/10.1016/j.bushor.2019.07.006

[9] Wirth, R., & Hipp, J. (2000, April). CRISP-DM: Towards a standard process model for data mining. In *Proceedings of the*

4th international conference on the practical applications of knowledge discovery and data mining (Vol. 1, pp. 29-39).

[10] Brzozowska, J., Pizoń, J., Baytikenova, G., Gola, A., Zakimova, A., & Piotrowska, K. (2023). Data engineering in CRISP-DM process production data–case study. *Applied Computer Science*, *19*(3). https://doi.org/10.35784/acs-2023-26

[11] El-Morr, C., Jammal, M., Ali-Hassan, H., & El-Hallak, W. (2022). Machine Learning for Practical Decision Making. *International Series in Operations Research and Management Science*. https://doi.org/10.1007/978-3-031-16990-8_4

[12] Pohl, M., Haertel, C., Staegemann, D., & Turowski, K. (2023). The Linkage to Business Goals in Data Science Projects. Retrieved from https://aisel.aisnet.org/acis2023/60/

[13] Ramadhan, Z. I., & Widiputra, H. (2024). Comparative Analysis of Recurrent Neural Network Models Performance in Predicting Bitcoin Prices. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, *8*(3), 377-388. https://doi.org/10.29207/resti.v8i3.5810

[14] Papathanasiou, J., Ploskas, N., Papathanasiou, J., & Ploskas, N. (2018). *Topsis* (pp. 1-30). Springer International Publishing. https://doi.org/10.1007/978-3-319-91648-4_1

[15] Chakraborty, S. (2022). TOPSIS and Modified TOPSIS: A comparative analysis. *Decision Analytics Journal*, *2*, 100021. https://doi.org/10.1016/j.dajour.2021.100021

[16] Hwang, C.-L., & Yoon, K. (1981). Multiple Attribute Decision Making. In Lecture Notes in Economics and Mathematical Systems. *Springer Berlin Heidelberg*. https://doi.org/10.1007/978-3-642-48318-9

[17] Cortes, C. (1995). Support-Vector Networks. *Machine Learning*

[18] Cheng, D., Zhang, S., Deng, Z., Zhu, Y., & Zong, M. (2014). k NN algorithm with data-driven k value. In *Advanced Data Mining and Applications: 10th International Conference, ADMA 2014, Guilin, China, December 19-21, 2014. Proceedings 10* (pp. 499-512). Springer International Publishing. https://doi.org/10.1007/978-3-319-14717-8_39

[19] Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5-32.

[20] Zhang, J., Liu, H., Bai, W., & Li, X. (2024). A hybrid approach of wavelet transform, ARIMA and LSTM model for the share price index futures forecasting. *The North American Journal of Economics and Finance*, *69*, 102022. https://doi.org/10.1016/j.najef.2023.102022

[21] Zaini, N., Ahmed, A. N., Ean, L. W., Chow, M. F., & Malek, M. A. (2023). Forecasting of fine particulate matter based on LSTM and optimization algorithm. *Journal of Cleaner Production*, *427*, *139233*. https://doi.org/10.1016/j.jclepro.2023.139233

[22] Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature. *Geoscientific model development*, *7*(3), 1247-1250. https://doi.org/10.5194/gmd-7-1247-2014

[23] Garcia-Arteaga, J., Zambrano-Zambrano, J., Parraga-Alava, J., & Rodas-Silva, J. (2024). An effective approach for identifying keywords as high-quality filters to get emergency-implicated Twitter Spanish data. *Computer Speech & Language*, *84*, 101579. https://doi.org/10.1016/j.csl.2023.101579

[24] Greenland, A. (2021). Deployment and Life Cycle Management. In *INFORMS Analytics Body of Knowledge* (pp. 275-310). INFORMS. https://doi.org/10.1002/9781119505914.ch8

[25] Shekhovtsov, A., & Kołodziejczyk, J. (2020). Do distance-based multi-criteria decision analysis methods create similar rankings?. *Procedia Computer Science*, *176*, 3718-3729. https://doi.org/10.1016/j.procs.2020.09.015