

DEMAND FORECASTING FOR IMPROVED INVENTORY MANAGEMENT IN SMALL AND MEDIUM-SIZED BUSINESSES

Dian Indri Purnamasari¹, Vynska Amalia Permadi^{2*}, Asep Saepudin³,
Riza Prapascatama Agusdin⁴

¹ Department of Accounting, Universitas Pembangunan Nasional Veteran Yogyakarta

^{2*,4} Department of Informatics, Universitas Pembangunan Nasional Veteran
Yogyakarta

³ Department of International Relations, Universitas Pembangunan Nasional Veteran
Yogyakarta

email: dian_indri@upnyk.ac.id¹, vynspermadi@upnyk.ac.id^{2*}, asep.saepudin@upnyk.ac.id³,
rizapra@upnyk.ac.id⁴

Abstract

Small and medium-sized businesses are constantly seeking new methods to increase productivity across all service areas in response to increasing consumer demand. Research has shown that inventory management significantly affects regular operations, particularly in providing the best customer relationship management (CRM) service. Demand forecasting is a popular inventory management solution that many businesses are interested in because of its impact on day-to-day operations. However, no single forecasting approach outperforms under all scenarios, so examining the data and its properties first is necessary for modeling the most accurate forecasts. This study provides a preliminary comparative analysis of three different machine learning approaches and two classic projection methods for demand forecasting in small and medium-sized leathercraft businesses. First, using K-means clustering, we attempted to group products into three clusters based on the similarity of product characteristics, using the elbow method's hyperparameter tuning. This step was conducted to summarize the data and represent various products into several categories obtained from the clustering results. Our findings show that machine learning algorithms outperform classic statistical approaches, particularly the ensemble learner XGB, which had the least RMSE and MAPE scores, at 55.77 and 41.18, respectively. In the future, these results can be utilized and tested against real-world business activities to help managers create precise inventory management strategies that can increase productivity across all service areas.

Keywords: Demand Forecasting, Inventory Management, Machine Learning, Small and Medium-Sized Businesses

Received: 09-01-2023 | Revised: 28-03-2023 | Accepted: 31-01-2023

DOI: <https://doi.org/10.23887/janapati.v12i1.57144>

INTRODUCTION

Data analytics is no longer limited to huge multinational companies due to advancements in data processing and storage capacity. However, it can be critical in developing strategies to support businesses of all sizes. On the other hand, small and medium-sized businesses should be able to use data analytics to uncover information such as client purchase habits,

demand forecasting, and effective customer relationship management. According to Bokman et al. [1], businesses that use consumer analytics beat their competition by 126%. In contrast, failing to consider data analysis in strategy formulation may put you at a competitive disadvantage.

Furthermore, increased consumer demand for goods and services has encouraged small and medium-sized

businesses to look for ways to improve operational efficiency in any service domain. Internal logistics and warehousing are high-growth sectors that can increase an organization's operational efficiency [2], [3] and have been the subject of several studies, as evidenced by the fact that influential organizations have conducted them [1]. Businesses have encouraged and supported the fulfillment of customer requirements and expectations throughout the supply chain and warehouse activities, which are critical to a company's operational performance. They discovered that inventory management significantly impacts logistics performance indices, particularly for organizations looking to reduce costs and improve product formulation and delivery procedures.

Inventory management refers to administering and maintaining a company's stock or inventory. In this context, the word "inventory" refers to raw materials, auxiliary materials, items in production, finished goods, and spare parts. Several intelligent inventory management solutions have been investigated to increase inventory management effectiveness [4]–[6]. Inventory management is closely linked to customer relationship management, and several technologies are used to meet consumer demand by shifting from traditional to innovative inventory management. Various approaches are available for forecasting client demand. Two possible strategies are to employ statistical analytic methodology and data mining approaches. A prediction based on statistics or mathematical analysis strongly relies on the quality of historical data (such as transaction history/product orders). As a result, if an organization has archives of high-quality data, making more accurate estimations will be easier. Organizations that use this inventory management technique typically have a quantitative analytic staff that analyzes historical data on a regular basis to uncover potential patterns and trends in market demand. The requirements analysis is based on historical customer demand transactions and will be used to manage the firm's inventories.

X. Guo's research [7] observed past order data to estimate client demand. The

initial investigation consisted primarily of scanning the company's website for information on product orders. Then, utilizing data mining tools, the study was undertaken on projected future consumer requests until a more efficient inventory management policy was established. Inventory management using data mining and sales history data may significantly reduce total inventory expenses. As previously stated, data quality is critical for achieving high-significance estimations and efficiency. However, the previous study showed that the gathered findings might still be improved/optimized because web-collected data is typically poor-quality and requires numerous additional procedures to be pre-processed. Another study [8] used data mining techniques to control inventories and focused on associating historical data with decision-making. A complete analysis was conducted based on the observed intercorrelation to boost forecast accuracy, and the authors believe that the company's expenses and consumption will be lowered. Moreover, a predictive analytics-based approach for inventory management that uses machine learning algorithms to analyze sales data and predict future demand has already been conducted. The authors focus on the use of time-series forecasting models to make accurate demand forecasts and optimize inventory levels. The proposed approach also incorporates a dynamic pricing model that adjusts prices based on demand patterns, which can help to further optimize inventory and reduce costs. The paper presents a case study demonstrating the proposed approach's effectiveness in improving inventory management for an online retailer.

Moreover, Zhang's [9] literature review comprehensively reviews the existing literature on inventory management for spare parts. The author discusses various approaches and techniques for spare parts inventory management, including forecasting demand, setting inventory levels, and determining reorder points. The paper also examines the challenges and opportunities associated with spare parts inventory management, such as intermittent demand, obsolescence, and service level

agreements. Overall, the paper provides a useful resource for researchers and practitioners seeking to improve spare parts inventory management in various industries, including manufacturing, aerospace, and healthcare.

This research aims to utilize various analysis techniques to determine the most appropriate model for one of the Small and Medium-Sized Leathercraft Businesses in Yogyakarta, Indonesia, as in previous studies. The selected leathercraft industry is among the largest SMEs in Yogyakarta, established in 2011, and sells its products through various channels. The company has not yet employed a data analytics approach in its decision-making process. Thus, this study seeks to provide valuable insights into the business by suggesting a demand forecasting model that aligns with service requirements, such as avoiding stock-outs or over-orders. To achieve the company's goal of delivering the best services to its customers, we will identify the most suitable demand forecasting technique that allows them to predict future monthly demand forecasts.

Since our objective is to develop the most relevant forecasting model concepts to assist the company in managing its inventory of handcrafted products across multiple parameters, the main takeaway of our study is that it is crucial to construct the forecasting model with particular consideration for the type of dataset in order to achieve higher accuracy. We will use a transactional historical time-series dataset to compare several demand forecasting algorithms. This work will make several contributions, including: (1) employing machine learning clustering to categorize various products with similar characteristics for efficiency, (2) utilizing demand forecasting based on statistical computation and machine learning algorithms to predict demand and provide a comparative analysis of each model's performance, and (3) suggesting the best forecasting model for one of the largest leathercraft SMEs in Yogyakarta's inventory management.

METHOD

Forecasting models can be classified into two types: qualitative and quantitative, each with three approaches based on analytical methodology: statistical, data mining/machine learning-based, and hybrid. The literature on business forecasting outlines continuous and intermittent demand techniques based on underlying market demands for products to manage inventory. Identifying trends such as intermittency is critical for selecting the best forecasting strategy. A typical approach that emphasizes expert judgment or client viewpoints above quantitative analysis is qualitative forecasting, also known as judgmental forecasting. Qualitative forecasting is desirable and frequently necessary in the absence of historical data to support quantitative methodologies or when past values have little or no influence on future events. However, qualitative forecasting is prone to bias due to its reliance on human opinion, which can be influenced by personal and political objectives. Experts and forecasters also tend to place greater emphasis on recent historical happenings, resulting in estimates that are close to the current reference point, adding another challenge to appraising predictions.

Quantitative forecasting uses mathematical (statistical) models to predict the future. Since quantitative forecasting models are objective, they should be used whenever there is a substantial amount of previous data that can be logically associated with predicted values (i.e., past historical data have unique trends and continuous values). Cross-sectional or time-series data can be used for quantitative forecasting, with time-series data being the most common type used for predictions. Quantitative forecasting uses various models based on unique combinations of predictive parameters and are included in one of two categories: time-series or explanatory. Explanatory forecasting models aim to identify the factors that affect the target variable, such as inflation, while not taking previous trends into consideration. Regression analysis is the most well-known method in the field of forecasting. Regression projections explore the

relationships between one or more dependent variables and an independent variable.

The following explanation covers a variety of time forecasting methodologies and current publications on demand forecasting applications, split into three categories: statistical, machine learning, and hybrid.

Statistical Procedures

Time-series forecasting, particularly demand and sales forecasting, is often accomplished through statistical methods. Furthermore, statistical methods can be broadly classified as either continuous, where a constant time-series pattern captures the demand history, or non-continuous. Additionally, there is an ad hoc intermittent special analytic technique for slow-moving items. As part of our literature review, in this section, we will comprehensively discuss several examples of well-known classic statistical forecasting approaches and some of the straightforward methodologies that are frequently used.

The implementation of forecasting methods on time series data relies mainly on Exponential Smoothing models. This method generates projections as the weighted average of previous data. There are several approaches to exponential smoothing, including simple, double, and triple Exponential Smoothing models. The Autoregressive Integrated Moving Average model [10] combines the best aspects of autoregressive and moving average models by differentiating time series data. ARIMA iteratively executes three stages using the Box-Jenkins method: model detection, parameter estimation, and diagnostic checking. The Theta Model [11] is frequently applied in time series forecasting, particularly in supply chain management and planning, because of the precision of its point forecasts [12]. The Vector Autoregressive (VAR) model [13] extends the univariate autoregressive model's capabilities to multivariate time series prediction. It has become a common tool for time series forecasting due to its ease of use and versatility. However, selecting the variables and lags employed in a VAR model is

essential. To keep this method performing well, limit the number of variables to the correlated ones.

Machine Learning Methods

Machine learning (ML) approaches were first used for forecasting in 1964, but little follow-up work was done for several decades. Since then, several studies have been conducted on applying ML systems to demand to forecast. The most popular time-series forecasting models include CART regression trees [14], Generalized Regression Neural Networks [15], K-Nearest Neighbor regression [16], Bayesian Neural Networks [17], Gaussian Processes regression [18], Long Short Term Memory network [19], Multi-Layer Perceptron [20], Recurrent Neural Networks [21], and Radial Basis Functions [22]. Several large-scale comparative studies, most of which are empirical, have reviewed different approaches to regression or time-series forecasting challenges. In a significant comparative study [23], the Multi-Layer Perceptron and Gaussian Processes regression were the most effective algorithms, followed by Bayesian Neural Networks and Support Vector Regression. Another study [24] using time-series data points discovered that the model with the best performance was Radial Basis Functions, followed by Recurrent Neural Networks and Multi-Layer Perceptron. Generalized Regression Neural Networks, on the other hand, had the worst performance according to recent comparison research [25], and Multi-Layer Perceptron is the most viable forecasting strategy among Machine Learning models. Although each method has advantages and weaknesses, data quality is critical for any empirical study seeking to evaluate the performance of a specific forecasting methodology. As such, there is no "generic" guaranteed technique for making predictions. The nature of the data and the situation in which the forecast is being made should affect the methodology chosen. Research also shows that Neural Networks and their variants perform the best of all machine learning algorithms when it comes to predicting time series.

Hybrid Approaches

This approach aims to bring together the best features of many statistical and ML-based forecasting techniques. Hybrid approaches include methods like SOM-SVR and ANN-ARIMA. To provide better learning and more precise prediction results, SOM-SVR uses a Self-Organizing Map to first split the dataset into clusters, then a Support Vector Regressor delivers grouped data. Tay and Caos' research [26] utilized the hybrid SOM-SVR for financial time series forecasting. Moreover, ANN-ARIMA has produced more accurate forecasts because of its ability to cover linear and nonlinear time series components [27]. Time-series forecasting problems, including the prediction of energy prices and stock market movements [28], have benefited from this hybrid approach.

We concluded from our thorough investigation that no single forecasting approach outperforms all scenarios. For the most accurate forecasts, it is necessary to first examine the data and its properties. This research aimed to evaluate which forecasting methodologies would be the most efficient for one of the leathercraft Small and Medium-Sized Businesses in Yogyakarta, Indonesia. The steps we took to complete our contribution are detailed in the following parts.

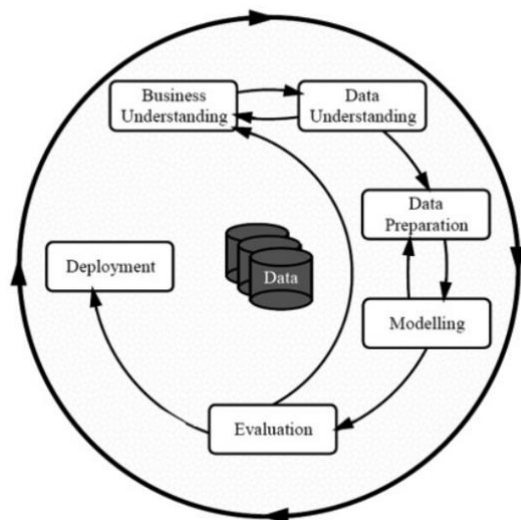


Figure 1. CRISP-DM methodology [29]

First, we applied a Cross Industry Standard Process for Data Mining (CRISP-DM) [29] approach to resolve our demand

forecasting comparative research. The stages of CRISP-DM are depicted in Fig. 1. CRISP-DM methodology is the most widely disseminated standard process model that describes common data mining techniques and phases. The initial phase is to analyze the business process and available data, followed by data preparation and modeling, and lastly, evaluate the model. This study, however, did not encompass the deployment step. The procedures involved at each level are outlined below.

Business & Data Understanding

After collecting the necessary transactional data from various sales channels in the leathercraft industry, we first evaluate the data that will be used to create a solution as the initial step in developing a demand forecast. Therefore, the exploratory data analysis process involves the following steps: first, visualizing historical sales and demand data; second, analyzing the primary characteristics of demand and sales time series; and third, examining the cross-correlation between demand and other time-dependent variables (such as product price). Additionally, we have incorporated certain features into the datasets. For example, we included vacation dates in our analysis to determine if external factors could aid in demand forecasting. While compiling holiday information, we considered both the general public's interests and observable school holiday periods.

Data Preparation

After completing the above stages, and considering that numerous products have already been sold, we recommend grouping the products into specific clusters to generate a demand prediction forecast for each group. To classify products with similar characteristics into multiple categories, we will use K-Means, a machine learning clustering approach. We will utilize hyperparameter tuning techniques such as the elbow method to determine the optimal number of k-clusters that reflect the best number of product groups. We do not wish to specify the exact number of categories required, as we aim to use a machine learning implementation to find the most

suitable number of clusters based on their characteristics.

Our objective here is to identify the optimal number of clusters that exhibit similar characteristics in the data. The parameter "k" is used to define the number of groups into which the data should be divided. The elbow method is utilized as a hyperparameter tuning technique to determine the ideal number of clusters. The appropriate number of clusters is achieved when the SSE sharply declines initially and then levels off as k increases. This phenomenon can be evaluated through the SSE plot of each k iteration.

Modeling

Table 1 summarizes the models investigated in this comparative analysis, aimed at comparing machine learning and traditional (classical) forecasting methods. Therefore, we chose SARIMA, a common univariate methodology, as well as SARIMAX, a multivariate classical alternative, to be compared with machine learning algorithms in our study. We also investigated machine learning approaches capable of capturing nonlinear relationships, taking into account the probability of temporal dependencies. For this purpose, we built a Recurrent Neural Network (RNN) using Long Short-Term Memory (LSTM) network cells. Additionally, we included Extreme Gradient Boosting in this comparative study, as integrated or ensemble approaches have been found to perform better than standalone ones in certain cases [30], [31]. Lastly, we included a Bayesian approach to model uncertainty using Gaussian Process Regression.

Table 1. Comparative analysis algorithm list
Algorithm Abbreviation

Algorithm	Abbreviation
Seasonal Autoregressive Integrated Moving Average	SARIMA
Seasonal Autoregressive Integrated Moving Average with external factors	SARIMAX
Long Short-Term Memory Network	LSTM
Extreme Gradient Boosting (XGBoost)	XGB
Gaussian Process Regression	GPR

We used time series cross-validation with regular model fitting to simulate the operational demand forecasting scenario with frequently updated business sales reports and a potentially shifting data distribution. Firstly, the 30 most recent months were selected as the training set to determine the best hyperparameter combination. The remaining data was used as the testing set. The model configuration with the best performance across the entire test set was chosen based on the evaluation metrics.

Evaluation

We used the Root Mean Squared Error (RMSE) in Equation 1 and the Mean Absolute Percent Error (MAPE) in Equation 2 for model evaluation. A lower score indicates improved performance across all assessment metrics. We added a constant value to the denominator to avoid division by zero when using MAPE. MAPE values can potentially increase significantly when demand numbers are close to zero, which is a common situation in daily observation datasets.

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

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (2)$$

where n is the number of data points, y_i is the i -th measurement, and \hat{y}_i is its corresponding prediction.

RESULT AND DISCUSSION

This section summarizes our findings, followed by an in-depth analysis of our experimental research. After studying the characteristics of historical sales data, we selected several product characteristics (Product Size, Product Motif, Product Color), sales amount, sales type, and unit pricing as clustering features. An overview of the data, including the attributes and value distributions for each attribute, is provided in Table 2.

Then, we employed K-means clustering to automatically cluster the data based on similar characteristics. Our aim was to determine the best number of clusters that represent the best data group with similar features. The k parameter indicates the number of groups into which the data should be aggregated. This research applies the elbow approach as hyperparameter tuning to determine the best number of clusters. The SSE is the elbow method's core index, representing the clustering impact indicating the clustering inaccuracy of all samples under evaluation of different k possibilities. When the number of clusters is fewer than k, increasing k amplifies the degree of aggregation of each cluster, hence reducing the SSE. k is considered as the correct amount of clusters when the SSE initially decreases sharply and then levels off as k increases.

Table 2. Data Overview

Attributes	Value Distributions
Product Name	More than thousand unique product name
Product Size	Small, Medium, Large
Product Motif	More than fifty unique product motif (batik motif, e.g.: Sidomukti, Parang, etc.)
Product Color	More than thirty unique product color (e.g. Havana, Borduk, etc.)
Sales Amount	12 ~ 200
Sales Type	In stock, Pre-order
Unit Pricing	75.000 ~ 1.195.000

For each iteration, test results to determine the proper number of k clusters and their SSE value are represented in the k and SSE diagram in Fig 2. We aimed to vary the value of k from 1 to 8, and the graph below shows that three clusters are sufficient for grouping our transactional dataset. Three is chosen since SSE initially decreases sharply as seen in the figure below, then levels off as k increases.

Then, the obtained k is used to group all the products into three cluster using K-Means algorithm. Through cluster result data analysis, we concluded that the third category of products has the highest sales volume, the highest sales frequency, and the lowest price. We labeled this category as "Prioritized product" since these products

have the highest customer demand and must always be kept in stock. The first category of products ranks second in terms of sales volume and frequency, and the unit price is also lower. So we named this group "Popular product," which should be kept in stock in moderate quantities. The second category of products is more expensive than the first and third categories, and its sales are still less frequent; hence we labeled these products as an "Exclusive Collection," which should be kept in limited quantities to provide alternative options while not interrupting cash flow. These classification results can be used to analyze whether the business's stock inventory is adequate on an operational basis.

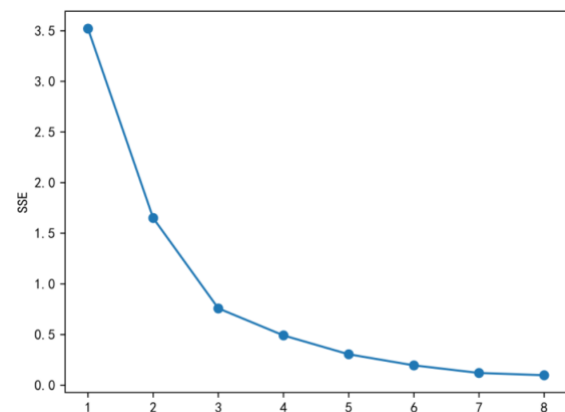


Figure 2. Elbow method SSE plot results to determine the best k cluster is three

The next step to achieve successful inventory management is to establish the exact amount of stock levels for each category mentioned above. The most straightforward strategy is to guess the proportion of stock in each category by trial and error. Unfortunately, such approaches impair a company's management team's ability to recognize critical moments. Therefore, the following section explores several demand forecasting methodologies to provide the manager with the best demand forecasting method, completing the inventory management process. Our analysis will only provide a rough estimate of which algorithm is most likely to produce the best results in our case study. In the future, the manager will need to implement a system that utilizes one of these demand forecasting methods to be

adequately prepared to develop inventory management strategies.

Before machine learning algorithms can use the series data, which contains historical transactional data from the last three years (2020-2022), it must be turned into a supervised learning scenario. In other words, we convert numerical inputs into outputs by taking the features of earlier time steps t_2 and t_1 as input and output the supervised value of the current time step t . Then, we divide the dataset into training and testing data. We use the first 30 months of data for training and the last 6 months for testing.

Hyperparameter tuning is a machine-learning technique that takes a snapshot of a model's performance at a given time and compares it to earlier snapshots. Every machine learning method requires the establishment of hyperparameters before developing a model into a real intelligence system. By adjusting the hyperparameters of the model, we can boost our model's performance and validate the chosen parameters using the validation dataset.

Before we describe our procedure further, it is important to understand the concept of cross-validation, which is a critical step in the hyperparameter tuning process. Cross-validation (CV) is a statistical technique used to evaluate the efficacy of machine learning models. Normally, we can predict how well a model will do on unknown data only after it has been trained. In other words, we cannot tell if the model is underfitting, overfitting, or performing well by evaluating how it performs on unknown data. When the data supplied is limited, cross-validation is considered a highly beneficial way to establish a machine learning model's performance. To perform cross-validation, we set aside some of the data for testing and validation, meaning not all subsets will be used to train the model; several data points will be kept for future use to validate the model's performance. K-Fold is a popular cross-validation strategy, and we used it to validate our model. We used a value of 10 for K in our cross-validation process for each algorithm. Tables 3–5 display the hyperparameter scenarios used by SARIMAX, LSTM, and GPR in this study.

Table 3. Seasonal Autoregressive Integrated Moving Average (with external factors) Hyperparameter

Parameter	Values
$p/q/P/Q$	[0, 1, 2, 3]
d/D	[0, 1]
$exog$	[False, True]
$transf$	[False, 'log', 'pw']

Table 4. Long Short-Term Memory Network Hyperparameter

Parameter	Values
activation_function	ReLU
optimizer	Adam
dropout_rate	[0.0, 0.5]
batch_size	[4, 8, 16, 32]
learning_rate	[1e-3, 1e-2, 1e-1]
min_val_loss_improvement	[0.01, 0.1]
max_epochs_wo_improvement	[10, 30, 50, 100]
lstm_hidden_dim	[5, 10, 15, 50]
lstm_num_layers	[1, 2]
seq_length	[1, 4]

Table 5. Extreme Gradient Boosting Hyperparameter

Parameter	Values
learning_rate	[0.05, 0.1, 0.3]
max_depth	[3, 5, 7]
subsample	[0.3, 0.7, 1]
n_estimators	[10, 100, 500, 1000]
gamma	[0, 1, 10]
alpha	[0, 0.1, 1, 10]
reg_lambda	[0, 0.1, 1, 10]

Below is a summary outlining several hyperparameters that can be adjusted while optimizing GPR. Since the kernel function is the most influential GPR parameter, we experimented with a wide variety of kernel functions as well as specific combinations of a maximum of three different kernels.

1. *SquaredExponential()*
2. *Matern52()*
3. *White()*
4. *RationalQuadratic()*
5. *Polynomial()*
6. *Periodic(kernels = SquaredExponential(), period = seasonal_periods)*
7. *Periodic(kernels = Matern52(), period = seasonal_periods)*
8. *Periodic(kernels = RationalQuadratic(), period = seasonal_periods)*

To thoroughly examine our experiments, we have prepared a summary of the results for each of the five demand

forecast methods investigated and the best evaluation metric results for each hyperparameter iteration. These results are shown in Table 6. SARIMA, the classic projection method, has an RMSE of 93.13 and a MAPE of 85.29. The SARIMAX algorithm, an extension of SARIMA, performs better with an RMSE of 48.96 but has a higher MAPE of 89.85. The LSTM algorithm has a RMSE of 85.36 and a MAPE of 82.41. The ensemble learner XGB outperforms other algorithms with a RMSE of 55.77 and a MAPE of 41.18, making it the most accurate algorithm for demand forecasting in this study. Finally, the Gaussian process regression (GPR) has a RMSE of 63.41 and a MAPE of 77.69.

Table 6. Comparison of metric evaluation results of all demand forecast model

Algorithm	Evaluation Metric	
	RMSE	MAPE
SARIMA	93.13	85.29
SARIMAX	48.96	89.85
LSTM	85.36	82.41
XGB	55.77	41.18
GPR	63.41	77.69

Our findings indicate that the machine learning-based techniques outperform the classic statistical approach in forecasting demand. XGB yields the best performance in terms of RMSE and MAPE, as highlighted in the table. In addition, the XGB ensemble outperformed the rest of the machine learning models. These findings demonstrate that the ensemble learning approach is preferable for capturing data phenomena in this use case. Compared to the other models evaluated, the results also reveal that XGB's superior performance can overcome the limited quantity of data used in this study, where only data from the previous three years were used. LSTM and GPR may not perform well in this case because the amount of information available significantly impacts the machine learning algorithms' performance. SARIMAX's comparable strong performance might be due to extrinsic variables such as distinct data dissemination in the training set, which could inhibit prediction. However, since all superior approaches were multivariate, the findings show that external inputs, such as holiday

data and clusters' generated characteristics, lead to increased predicting ability. This extra information is also considered valuable in machine learning-based approaches.

Based on the evaluation matrix, we will recommend the use of XGB projection as the demand forecasting model for our case study, one of Yogyakarta's Leathercraft Small and Medium-Sized Businesses. However, it is crucial to consider the impact of the new normal era after COVID-19. We suggest first using historical data as training data to determine if it fits well, as there may be different sales trends during the pandemic.

CONCLUSION

This study presents a preliminary comparative analysis of three Machine Learning approaches and two classic projection methods for demand forecasting in Leathercraft Small and Medium-Sized Businesses. Firstly, we utilized K-means clustering to group the products into three clusters based on the similarity of product characteristics, using the elbow method's hyperparameter tuning. Furthermore, the data was summarized to represent the significance of the clustering results. The "Prioritized product" cluster had the highest sales volume and frequency, and the price was also relatively low. The products in the "Popular product" cluster ranked second in terms of sales volume and frequency, and the price was also relatively low. The "Exclusive collection" cluster contained more expensive products than the first two categories, and its sales were less frequent, and the unit price was also lower. According to strategic inventory management recommendations, "Popular Products" should be retained in moderation, "Priority Products" should be kept in large quantities, and "Exclusive Collections" should be preserved in limited quantities to provide alternative options while not interrupting cash flow.

The evaluation results of five different demand forecasting algorithms, including a classic projection method, are presented. In the evaluation of demand forecasting algorithms, SARIMA, which is a traditional projection method, produces a RMSE of

93.13 and a MAPE of 85.29. On the other hand, the SARIMAX algorithm, which is a modification of SARIMA, performs better with an RMSE of 48.96, but its MAPE is higher at 89.85. In terms of the LSTM algorithm, it generates a RMSE of 85.36 and a MAPE of 82.41. In comparison, the ensemble learner XGB stands out as the most accurate algorithm for demand forecasting in this study, with a RMSE of 55.77 and a MAPE of 41.18. Lastly, the Gaussian process regression (GPR) has a RMSE of 63.41 and a MAPE of 77.69. Overall, the results show that machine learning algorithms outperform the classic projection method, and XGB is the most accurate algorithm for demand forecasting in this study.

Furthermore, in the future, these results can be utilized and tested against real-world business activities to help managers create accurate inventory management strategies. As a suggestion for subsequent implementation, considering the new normal era after COVID-19, historical data should be used as training data first to determine if it fits well since there may be a different trend during the pandemic in terms of sales data.

REFERENCES

- [1] McKinsey & Company, "Using customer analytics to boost corporate performance Marketing Practice Key insights from McKinsey's DataMatics 2013 survey," 2014.
- [2] A. Singh, M. Wiktorsson, and J. B. Hauge, "Trends In Machine Learning To Solve Problems In Logistics," *Procedia CIRP*, vol. 103, pp. 67–72, 2021, doi: 10.1016/j.procir.2021.10.010.
- [3] A. Tufano, R. Accorsi, and R. Manzini, "A machine learning approach for predictive warehouse design," *The International Journal of Advanced Manufacturing Technology*, vol. 119, no. 3–4, pp. 2369–2392, Mar. 2022, doi: 10.1007/s00170-021-08035-w.
- [4] A. M. Atieh *et al.*, "Performance Improvement of Inventory Management System Processes by an Automated Warehouse Management System," *Procedia CIRP*, vol. 41, pp. 568–572, 2016, doi: 10.1016/j.procir.2015.12.122.
- [5] T. Lima de Souza, D. Barbosa de Alencar, A. P. Tregue Costa, and M. C. Aparício de Souza, "Proposal for Implementation of a Kanban System in the Auxiliary Inventory Sector in a Auto Parts Company," *Int J Innov Educ Res*, vol. 7, no. 10, pp. 849–859, Oct. 2019, doi: 10.31686/ijer.vol7.iss10.1833.
- [6] Xie Haiyan, "EMPIRICAL STUDY OF AN AUTOMATED INVENTORY MANAGEMENT SYSTEM WITH BAYESIAN INFERENCE ALGORITHM," *Int J Res Eng Technol*, vol. 04, no. 10, pp. 398–405, Oct. 2015, doi: 10.15623/ijret.2015.0410065.
- [7] X. Guo, C. Liu, W. Xu, H. Yuan, and M. Wang, "A Prediction-Based Inventory Optimization Using Data Mining Models," in *2014 Seventh International Joint Conference on Computational Sciences and Optimization*, Jul. 2014, pp. 611–615. doi: 10.1109/CSO.2014.118.
- [8] L. Lin, W. Xuejun, H. Xiu, W. Guangchao, and S. Yong, "Enterprise Lean Catering Material Management Information System Based on Sequence Pattern Data Mining," in *2018 IEEE 4th International Conference on Computer and Communications (ICCC)*, Dec. 2018, pp. 1757–1761. doi: 10.1109/CompComm.2018.8780656.
- [9] S. Zhang, K. Huang, and Y. Yuan, "Spare Parts Inventory Management: A Literature Review," *Sustainability*, vol. 13, no. 5, p. 2460, Feb. 2021, doi: 10.3390/su13052460.
- [10] Box George E. P., Jenkins Gwilym M., Reinsel Gregory C., and Ljung Greta M., *Time Series Analysis: Forecasting and Control*. Wiley, 2015.
- [11] V. Assimakopoulos and K. Nikolopoulos, "The theta model: a decomposition approach to forecasting," *Int J Forecast*, vol. 16, no. 4, pp. 521–530, Oct. 2000, doi: 10.1016/S0169-2070(00)00066-2.
- [12] K. Nikolopoulos, V. Assimakopoulos, N. Bougioukos, A. Litsa, and F. Petropoulos, "The Theta Model: An Essential Forecasting Tool for Supply Chain Planning," 2011, pp. 431–437. doi: 10.1007/978-3-642-25646-2_56.
- [13] H. Lütkepohl, "Vector autoregressive models," in *Handbook of Research Methods and Applications in Empirical Macroeconomics*, Edward Elgar Publishing, 2013, pp. 139–164. doi: 10.4337/9780857931023.00012.
- [14] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification And Regression Trees*. Routledge, 2017. doi: 10.1201/9781315139470.

- [15] D. F. Specht, "A general regression neural network," *IEEE Trans Neural Netw*, vol. 2, no. 6, pp. 568–576, 1991, doi: 10.1109/72.97934.
- [16] N. S. Altman, "An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression," *Am Stat*, vol. 46, no. 3, p. 175, Aug. 1992, doi: 10.2307/2685209.
- [17] D. J. C. MacKay, "Bayesian neural networks and density networks," *Nucl Instrum Methods Phys Res A*, vol. 354, no. 1, pp. 73–80, Jan. 1995, doi: 10.1016/0168-9002(94)00931-7.
- [18] C. K. I. Williams and C. E. Rasmussen, "Gaussian Processes for Regression," in *Proceedings of the 8th International Conference on Neural Information Processing Systems*, 1995, pp. 514–520.
- [19] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.
- [20] Zhengyou Zhang, M. Lyons, M. Schuster, and S. Akamatsu, "Comparison between geometry-based and Gabor-wavelets-based facial expression recognition using multi-layer perceptron," in *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 454–459, doi: 10.1109/AFGR.1998.670990.
- [21] Medsker L.R. and Jain L.C., *Recurrent neural networks: design and applications*. CRC Press, 1999.
- [22] Buhmann M. D., *Radial Basis Functions: Theory and Implementations*, vol. 12. Cambridge University Press, 2003.
- [23] N. K. Ahmed, A. F. Atiya, N. el Gayar, and H. El-Shishiny, "An Empirical Comparison of Machine Learning Models for Time Series Forecasting," *Econom Rev*, vol. 29, no. 5–6, pp. 594–621, Aug. 2010, doi: 10.1080/07474938.2010.481556.
- [24] J. J. Montaña Moreno, A. Palmer Pol, and P. Muñoz Gracia, "Artificial neural networks applied to forecasting time series.," *Psicothema*, vol. 23, no. 2, pp. 322–9, Apr. 2011.
- [25] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "Statistical and Machine Learning forecasting methods: Concerns and ways forward," *PLoS One*, vol. 13, no. 3, p. e0194889, Mar. 2018, doi: 10.1371/journal.pone.0194889.
- [26] F. E. H. Tay and L. J. Cao, "Improved financial time series forecasting by combining Support Vector Machines with self-organizing feature map," *Intelligent Data Analysis*, vol. 5, no. 4, pp. 339–354, Nov. 2001, doi: 10.3233/IDA-2001-5405.
- [27] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, Jan. 2003, doi: 10.1016/S0925-2312(01)00702-0.
- [28] C. N. Babu and B. E. Reddy, "A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data," *Appl Soft Comput*, vol. 23, pp. 27–38, Oct. 2014, doi: 10.1016/j.asoc.2014.05.028.
- [29] R. Wirth and J. Hipp, "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*, 2000, pp. 29–39.
- [30] F. Petropoulos, R. J. Hyndman, and C. Bergmeir, "Exploring the sources of uncertainty: Why does bagging for time series forecasting work?," *Eur J Oper Res*, vol. 268, no. 2, pp. 545–554, Jul. 2018, doi: 10.1016/j.ejor.2018.01.045.
- [31] C. S. Bojer and J. P. Meldgaard, "Kaggle forecasting competitions: An overlooked learning opportunity," *Int J Forecast*, vol. 37, no. 2, pp. 587–603, Apr. 2021, doi: 10.1016/j.ijforecast.2020.07.007.