International Journal of Digital Entrepreneurship and Business (IDEB)
Vol. 7, No. 1, April 2026 E-ISSN 2723-5432
pp. 61 – 80 https://ejournal.uniji.ac.id/ideb/

Digital Business Transformation Through XGBoost-Based Sales Forecasting for Fast Food Franchise Inventory

Meinarini Catur Utami[1*], Mim Hanifah Permana[2], Elvi Fetrina[3]
[1,2,3]*Department of Information Systems, Faculty of Science and Technology,*
UIN Syarif Hidayatullah, Jakarta, Indonesia
**Corresponding Email: meinarini@uinjkt.ac.id*

ABSTRACT

Fast food franchises increasingly rely on data-driven decision-making in inventory planning, yet an Indonesian franchise with more than 300 outlets lacked an accurate sales forecasting system; overstocking 30 products per day, along with recurring understocking, triggered negative customer reviews that threatened business sustainability. This study developed a digital forecasting model to estimate daily sales, safety stock levels, and reorder points for two outlets, referred to as Outlet A and Outlet B, to preserve business confidentiality. Daily sales records from January 2023 to June 2024, combined with national holiday data, were processed using the extreme gradient boosting algorithm in Python and evaluated with root mean squared error and mean absolute percentage error; safety stock and reorder points were then calculated in Microsoft Excel. The model predicted an average of 435 products per day at Outlet A, with an error of 61.89 units (10.2 percent), and 360 products at Outlet B, with an error of 56.20 units (14.1 percent); both were categorized as good forecasts. Safety stock and reorder point values reached 98 and 968 units for Outlet A and 120 and 840 units for Outlet B, showing that the more volatile outlet required a proportionally larger buffer. The findings demonstrate that machine learning forecasting can be linked directly to classical inventory control within a single digital workflow, giving franchise businesses a practical route to reducing overstock and understock. This integration of outlet-level forecasting and inventory-control calculation in a perishable-product franchise setting in Indonesia constitutes the study's originality.

Keywords: *sales forecasting; XGBoost; digital business transformation; inventory efficiency; fast food franchise*

JEL Classification: *C53; M11; L81; O33*

ARTICLE INFO	
Article History Received: August 25, 2026 Revised: August 27, 2026 Accepted: September 04, 2026 Published Online: September 10, 2026 **How to cite** Utami, M. C., Permana, M. H., & Fetrina, E. (2026). Digital business transformation through XGBoost-based sales forecasting for fast food franchise inventory. International Journal of Digital Entrepreneurship and Business (IDEB), 7(1), 61–80.	**DOI** https://doi.org/10.52238/ideb.v7i1.416

1. Introduction

Consumption patterns have shifted globally toward fast food as consumers increasingly value efficiency in obtaining daily meals. The global fast-food market was valued at USD 772.04 billion in 2023 (Coherent Market Insight, 2026). Market slump during 2019-2020 due to pandemic social restrictions (Dairu & Shilong, 2021; Hasan, 2024). However, it recovered from 2020 to 2023 (Heaviside et al., 2020; Petropoulos et al., 2022) and is projected to grow at a CAGR of 4.89% from 2024 to 2032 (Allied Market Research, 2020). The sector therefore remains on the move, increasingly driven by digital technology and pressing operators to plan production and inventory with more precision.

The shift toward digitally enabled operations is particularly consequential for franchise businesses, which must coordinate production planning across a large and often geographically dispersed network of outlets. (Ekankumo, 2023). Because each outlet faces a distinct local demand pattern, franchise headquarters increasingly rely on digital tools rather than manual estimation to plan daily production quantities, allocate raw materials, and schedule replenishment. Digital transformation in this context is not limited to customer-facing technology such as mobile ordering; it extends to back-of-house decision support, including demand forecasting and inventory optimization, which directly affect cost efficiency and service quality. (Megawati et al., 2024).

This study focuses on a fast-growing Indonesian fast-food franchise that operates more than 300 outlets nationwide, referred to throughout this paper as “the Company” to preserve the confidentiality of the business and its internal data. The Company’s competitive edge relies heavily on steady outlet-level sales performance. Two of its outlets (Outlet A and Outlet B, anonymized here) recorded significant sales growth in 2023 compared to the previous year and are therefore suitable cases for studying the potential of digital forecasting technology to support operational decision-making at the outlet level.

The Company maintains strict operating standards emphasizing service and product quality, including a First In First Out (FIFO) inventory system to ensure product freshness: frozen chicken that has started thawing must be fried on the same day and withdrawn from inventory if unsold, and fried chicken must be sold within two hours of preparation. These standards are core to the Company’s brand promise. However, they also remove the error buffer in production planning, because any mismatch between forecasted and actual demand cannot be carried over to the next day as it might be for non-perishable goods, and is realized immediately as either wasted product or a stockout.

The Company regularly faces overstock situations for approximately 30 products per day and understock situations that generate negative customer reviews on digital platforms, due to the lack of an accurate, technology-enabled sales forecasting system (Huo, 2021; Husein et al., 2021). Both problems carry financial and reputational costs: overstock wastes raw materials and labor under the FIFO withdrawal policy, while understock damages customer experience at a time when digital word-of-mouth can amplify a single bad visit into a public review visible to prospective customers researching the brand online. In an increasingly digital market, these inefficiencies threaten business sustainability.

Previous studies show that machine learning methods such as Extreme Gradient Boosting (XGBoost) can improve sales prediction accuracy and support making inventory decisions in various sectors, including retail, e-commerce, and manufacturing. (Dairu & Shilong, 2021; Panarese et al., 2022; Turgut & Erdem, 2022). Other studies have addressed inventory control methods such as safety stock and reorder point calculation mostly separately from the forecasting process, which should, in principle, feed into them. (Heaviside et al., 2020; Nissa et al., 2023). What is missing is research that operationalizes machine-learning forecasting and classical inventory-control calculations as a single, integrated workflow within a franchise fast-food business in a developing-country context such as Indonesia, where perishability makes forecasting errors particularly costly.

This study addresses that gap with four objectives: to develop a daily sales forecasting model for two franchise outlets using the XGBoost algorithm; to evaluate the accuracy of the resulting forecasts using RMSE and MAPE; to calculate safety stock and reorder point values based on the forecast output; and to translate these technical results into practical digital-business recommendations that support the Company's broader digital transformation and organizational efficiency. In doing so, the study contributes both a methodological demonstration of integrating forecasting with inventory-control calculation, and a practical reference case for other franchise businesses facing similar overstock and understock challenges.

The remainder of this paper is organized as follows. The Literature Review discusses digital transformation in franchise business, sales forecasting theory, the XGBoost algorithm, forecast evaluation metrics, and inventory-control theory, and synthesizes prior empirical studies. The Method section describes the research design, data source, feature engineering, model development, and the inventory-control formulas. The Findings and Discussion sections present the forecasting outcomes, inventory-control values, digital-business recommendations, and the study's contributions, followed by the conclusion and limitations.

Indonesia is a particularly relevant setting. As one of the largest fast food markets in Southeast Asia, it has seen rapid growth in local and international franchise brands alongside rapid digital adoption among small and medium enterprises. (Megawati et al., 2024). However, much of this adoption has focused on customer-facing functions such as mobile ordering and payment. In contrast, back-end functions such as demand forecasting and inventory planning remain manual or dependent on the informal experience of outlet supervisors. This imbalance between front-end and back-end digitalization is the gap this study seeks to narrow, using a franchise fast food case as a practically grounded example.

Digital entrepreneurship scholarship increasingly frames data-driven tools such as forecasting models and dashboards as a core capability differentiating resilient, growth-oriented small and franchise businesses from those that remain reliant on manual planning (Ekankumo, 2023; Megawati et al., 2024). The technical contribution of an accurate forecasting model is inseparable from its organizational contribution: a model that staff does not trust or use has limited practical value regardless of its statistical accuracy. This study therefore treats the technical and organizational dimensions of digital transformation as interdependent, rather than presenting the model in isolation from how Company staff would actually use it.

Consistent with the objectives above, this study is guided by four research questions: (RQ1) How accurately can an XGBoost-based model forecast daily sales at the outlet level for a fast-food franchise business, as measured by RMSE and MAPE? (RQ2) What safety stock and reorder point values result from applying standard inventory-control formulas to this forecast output? (RQ3) How do forecast accuracy and the resulting inventory-control parameters differ between outlets with different demand patterns? Moreover, (RQ4) What digital-business recommendations can translate these technical results into practical operational improvements for a franchise organization? These research questions structure the presentation of results and are revisited explicitly in the Conclusion.

2. Literature Review

2.1. Digital Transformation and Digital Entrepreneurship in Franchise Business

Digital transformation refers to integrating digital technology into business processes to change how a company operates and creates value for customers. It has been a significant enabler of long-term sustainability for small, medium, and franchise-based culinary businesses, allowing owners to take faster, more data-driven operational decisions instead of depending on intuition or manual record-keeping (Megawati et al., 2024). It sits within the broader field of digital entrepreneurship, which uses online tools for marketing, transaction processing, and, increasingly, backend decision support such as demand forecasting and inventory optimization.

Franchise businesses are a distinctive case for digital transformation because they must balance operational standardization across outlets with genuine differences in local demand. (Ekankumo, 2023). A forecasting and inventory-control system that is replicable across outlets yet produces outlet-specific parameters is therefore valuable: it preserves the consistency franchising requires while respecting local variation that a one-size-fits-all approach would miss.

2.2. Sales Forecasting in Business Operations

Sales forecasting estimates future sales volume from historical and contextual information. It underpins budgeting and operational planning across most business functions, including procurement, staffing, and production scheduling. (Husein et al., 2021). Traditionally, it has relied on statistical time-series methods such as moving averages and exponential smoothing, or qualitative methods such as expert opinion. (Petropoulos et al., 2022). They remain applicable when demand is stable, but perform poorly when demand is shaped by interacting factors such as day-of-week effects, national holidays, and promotional activity, as is the case for fast food retail sales.

Within digital business transformation, forecasting is increasingly carried out with machine learning rather than statistical methods alone, since these algorithms capture non-linear relationships between several explanatory variables and the sales outcome and can be retrained automatically as new data arrives. (Raizada & Saini, 2021). This shift is itself a form of digital transformation in the forecasting function, and it is the approach adopted in this study.

2.3. Machine Learning Approaches to Demand Forecasting

Machine learning algorithms such as linear regression, support vector machines, artificial neural networks, and tree-based ensembles have been applied to sales and demand forecasting. Comparative studies find that tree-based ensembles, particularly gradient boosting variants, achieve competitive or superior accuracy relative to single-model approaches across a variety of retail and e-commerce forecasting tasks. (Huo, 2021; Panarese et al., 2022). Ensemble methods are also attractive practically: they are more robust to noisy or incomplete data than deep learning approaches and do not require the very large training datasets deep networks need, a relevant consideration for a dataset spanning roughly a year and a half.

2.4. Comparison of Forecasting Paradigms: Statistical versus Machine Learning Methods

It is useful to situate XGBoost relative to the classical statistical methods it increasingly complements or replaces. Statistical methods such as moving averages, exponential smoothing, and ARIMA (Autoregressive Integrated Moving Average) rely primarily on past values of the series to be forecast, and work well when demand is stable and dominated by a few clear seasonal or trend patterns. (Petropoulos et al., 2022). Their advantages are simplicity, interpretability, and low data requirements; their key limitation is that they struggle to incorporate multiple exogenous variables, such as a holiday flag, without additional model extensions. Tree-based ensembles such as XGBoost, by contrast, incorporate mixed numerical and categorical features directly and learn non-linear interactions between them, such as day-of-week by holiday status, at the cost of reduced interpretability and greater need for careful validation against overfitting. Table 1 summarizes this comparison across the criteria most relevant to franchise inventory planning.

Table 1: Comparison of Statistical and Machine Learning Forecasting Approaches

Criterion	Statistical Methods (e.g., ARIMA, Exponential Smoothing)	Machine Learning Methods (e.g., XGBoost)
Handling of multiple features (e.g., holidays, day-of-week)	Limited; typically requires model extensions	Native; features are incorporated directly
Interpretability	High; parameters have direct statistical meaning	Moderate; requires feature-importance analysis
Data volume required	Low to moderate	Moderate; benefits from at least 1–2 years of daily data
Robustness to irregular demand spikes	Lower without manual adjustment	Higher, if spikes are represented in training data
Typical implementation effort	Lower	Higher, but manageable with modern libraries

Source: authors' compilation

Given the multiple demand drivers relevant to fast food sales (weekly seasonality, national holidays, and outlet-specific promotional patterns), and given that at least a year of daily transaction data was available for both outlets, this comparison supports the choice of XGBoost as the primary forecasting method for this study, consistent with its use in comparable retail forecasting research. (Dairu & Shilong, 2021; Turgut & Erdem, 2022).

2.5. Extreme Gradient Boosting (XGBoost)

Extreme Gradient Boosting (XGBoost) is a scalable tree-based ensemble learning algorithm introduced by (Chen & Guestrin, 2016) and has been widely used in classification and regression tasks, such as sales forecasting (Dairu & Shilong, 2021). It implements the gradient boosting framework, training a sequence of decision trees, each correcting the residual errors of the previous trees. The model's prediction for observation i after adding K trees can be expressed as an additive function of the form shown below, where each f_k represents an individual regression tree, and x_i denotes the feature vector for observation i.

$$\hat{y}_i = \Sigma_{k=1}^{K} f_k(x_i),\ f_k \in \mathscr{F}$$

Training minimizes a regularized objective function: a differentiable loss term measuring the discrepancy between predicted and actual sales, plus a regularization term that penalizes model complexity to avoid overfitting. This regularization distinguishes XGBoost from earlier gradient boosting implementations and helps explain why it has repeatedly outperformed conventional forecasting approaches in applied studies, from retail produce sales prediction (Turgut & Erdem, 2022) to fraud detection in health insurance claims (Nabrawi & Alanazi, 2023).

2.6. Forecast Evaluation: RMSE and MAPE

In this study, the model performance is evaluated with two complementary metrics. Root Mean Squared Error (RMSE) is the square root of the mean squared difference between predicted and actual sales; because its scale matches the target, it is easy to interpret the average forecasting error directly in terms of units sold. (Putra, 2019). RMSE is defined as follows, where y_i is the actual value, $\hat{y}_i$ is the predicted value, and n is the number of observations.

$$RMSE = \sqrt{((1/n)\ \Sigma_{i=1}^{n}\ (y_i - \hat{y}_i)^2)}$$

Mean Absolute Percentage Error (MAPE) expresses the average error as a percentage of the actual value, making it comparable across outlets or products with different sales volumes. (Raizada & Saini, 2021). MAPE is defined as follows.

$$MAPE = (1/n)\ \Sigma_{i=1}^{n}\ |(y_i - \hat{y}_i)\ /\ y_i| \times 100\%$$

Following common practice in the forecasting literature, a MAPE value between 10% and 20% is generally categorized as a good forecast, while a MAPE below 10% is categorized as highly accurate; this study adopts that classification when interpreting its results.

2.7. Inventory Management: Safety Stock and Reorder Point

Inventory refers to assets a company holds either for sale or for use in an ongoing production process. Effective inventory management affects organizational efficiency: too little inventory risks stockouts, while too much ties up capital and, for perishable products, increases waste (Aghnani & Perdana, 2019). Because demand fluctuates and is rarely predicted perfectly, companies maintain a buffer known as safety stock, intended to absorb forecast error and demand variability during the replenishment lead time (Heaviside et al., 2020). A common formulation expresses safety stock as a function of a desired service level, the standard deviation of demand, and the lead time, as shown below, where Z is the service-level factor, σ is the standard deviation of demand, and L is the lead time.

$$Safety\ Stock\ (SS) = Z \times \sigma \times \sqrt{L}$$

The reorder point (ROP) is the inventory level at which a new order must be placed so that stock does not run out before the replenishment arrives. It is calculated by combining expected demand during the lead time with the safety stock buffer, as shown below, where $\bar{d}$ is the average daily demand.

$$Reorder\ Point\ (ROP) = (\bar{d} \times L) + SS$$

These formulas show why an accurate forecast is a prerequisite for sound inventory control: both the average demand and standard deviation terms that drive safety stock and reorder point are outputs of the forecasting process, the link this study operationalizes for two franchise outlets. (Nissa et al., 2023; Wanti et al., 2020).

2.8. Time Series Characteristics of Retail Sales Data

Daily retail sales data, including fast food sales, typically exhibit several characteristic patterns that a forecasting method must be able to capture: weekly seasonality driven by weekday versus weekend consumer behavior, calendar effects around national holidays and long weekends, gradual trend components reflecting outlet growth or decline, and irregular shocks caused by local events, weather, or promotional activity. As summarized in Table 1, tree-based ensembles are better suited than classical time-series models to data of this kind, because they learn non-linear interactions such as day-of-week by holiday status without the researcher specifying them in advance — the main theoretical reason XGBoost fits a dataset in which sales are shaped jointly by weekly patterns and irregular holiday effects.

2.9. Digital Entrepreneurship and Organizational Efficiency

Organizational efficiency is a company’s ability to meet its operating goals while minimizing the waste of resources, such as inventory, labor, and sales opportunities. In the digital entrepreneurship literature, technology adoption is conceptualized as a driver of efficiency, replacing informal, experiential judgment that is inconsistent across employees and hard to scale across outlets, with systematic data-driven decision rules (Megawati et al., 2024). Sales forecasting and inventory-control automation, of the kind developed in this study, are a direct application of this lever: they replace an outlet supervisor's informal daily estimate of how much chicken to thaw with a statistically grounded, auditable, and repeatable calculation, which is more easily standardized and monitored across a franchise network than tacit individual judgment.

2.10. Customer Experience and Digital Word-of-Mouth

A further strand concerns how operational failures such as stockouts translate into customer-facing consequences in a digitally connected retail environment. Where dissatisfaction was once confined to word-of-mouth within a customer's immediate social circle, digital review platforms and social media now allow a single negative experience to be documented publicly and influence prospective customers with no direct connection to the original event. For a franchise brand with hundreds of outlets, an inventory failure at one outlet implies a reputational risk beyond the outlet's own customer base, highlighting the strategic importance of accurate demand forecasting and reliable stock availability beyond its operational significance (Huo, 2021; Husein et al., 2021). This link between backend inventory accuracy and front-end digital reputation motivates treating this study's technical contribution as a digital-business, rather than purely computational, contribution.

2.11. Digital Business and Franchise Growth in Emerging Markets

Franchise business models have grown rapidly in emerging markets such as Indonesia, where they scale a proven operating concept across a dispersed consumer base while distributing capital investment across franchisees. But growth also magnifies demand-planning challenges: a mistake that would affect a single independent restaurant is repeated in every outlet sharing the same standardized menu and food-safety procedures (Ekankumo, 2023). Digital adoption in many emerging-market franchise networks has been limited to customer-facing functions such as mobile ordering and loyalty programs. In contrast, backend functions such as demand forecasting and inventory control rely on manual, outlet-level judgment. (Megawati et al., 2024). This asymmetry motivates the present study, which digitalizes the backend planning function rather than the customer-facing side of the business.

Another characteristic of many emerging-market franchise settings is that historical sales data, while available, is often underutilized beyond basic bookkeeping and tax reporting. Where digital point-of-sale systems already record every transaction, as is increasingly the case even for small and medium fast-food outlets in Indonesia, this data is a low-cost, readily available input for machine learning-based forecasting and requires no additional data-collection investment. The marginal cost of adopting a forecasting-based inventory workflow is therefore primarily analytical and organizational rather than one of new hardware or data infrastructure.

2.12. Summary of Related Studies

Table 2 synthesizes the prior studies most relevant to this research by forecasting method, application domain, and reported accuracy where available. It shows that while XGBoost and related gradient boosting methods have been applied successfully in retail and adjacent domains, few studies combine the forecasting step directly with inventory-control calculations such as safety stock and reorder point within a single applied case, which is the gap this study addresses.

Table 2: Summary of Related Studies on Machine Learning-Based Forecasting and Inventory Control

Criterion	Statistical Methods (e.g., ARIMA, Exponential Smoothing)	Machine Learning Methods (e.g., XGBoost)	Key Finding
Dairu & Shilong (2021).	XGBoost	Retail sales forecasting	XGBoost outperformed baseline models in prediction accuracy
Huo (2021)	Machine learning (general)	E-commerce sales prediction	ML methods improved prediction accuracy over statistical baselines
Panarese et al. (2022)	Gradient boosting	Retail sales forecasting platform	Gradient boosting provided reliable short-term sales forecasts
Turgut & Erdem (2022).	XGBoost	Retail produce sales	XGBoost achieved competitive accuracy for perishable products

Nabrawi & Alanazi (2023).	SVM & XGBoost	Health insurance fraud detection	XGBoost outperformed SVM on structured tabular data
Nissa et al. (2023).	Q-model inventory control	MSME inventory management	Forecast-based inventory control reduced lost sales risk
Heaviside et al. (2020).	EOQ, forecast & ROP	Retail minimum stock determination	Integrating forecast with EOQ/ROP improved stock adequacy
Authors (2026)	XGBoost + Safety Stock/ROP	Fast food franchise inventory	Integrated forecasting–inventory workflow with dashboard proposal

Source: authors' compilation

2.13. Research Novelty and Positioning

Table 3 positions this study relative to the most closely related prior works discussed above, across four dimensions: whether the study uses a machine learning forecasting method, whether it explicitly calculates safety stock and/or reorder point from the forecast output, whether it addresses a franchise or multi-outlet business context, and whether it proposes an accompanying digital monitoring tool. As the table shows, no single prior study combines all four elements, which is the specific novelty this research contributes.

Table 3: Positioning of This Study Relative to Prior Research

Study	***ML Forecasting***	***Safety Stock/ROP Calculation***	***Franchise/ Multi-outlet Context***	***Digital Dashboard Proposed***
Dairu & Shilong (2021).	Yes	No	No	No
Panarese et al. (2022)	Yes	No	No	No
Turgut & Erdem (2022)	Yes	No	No	No
Nissa et al. (2023)	No	Yes	No	No
Heaviside et al. (2020)	No	Yes	No	No
This study (2026)	Yes	Yes	Yes	Yes

Source: authors' compilation

Building on the theoretical foundations above, this study adopts a conceptual framework in which historical sales and holiday data feed into an XGBoost forecasting model; the resulting forecasts, together with their standard deviation, feed into safety stock and reorder point calculations; and both the forecasts and the inventory-control values feed into a proposed digital monitoring dashboard that supports day-to-day managerial decision-making. Table 4 summarizes this framework stage by stage.

Table 4: Conceptual Framework Linking Forecasting, Inventory Control, and Digital Monitoring

Stage	***Input***	***Process***	***Output***
1. Forecasting	Historical sales, holiday data	XGBoost model training and validation	Daily sales forecast per outlet
2. Inventory Control	Sales forecast, standard deviation	Safety stock and reorder point calculation	Safety stock and reorder point values
3. Digital Monitoring	Forecast and inventory-control output	Dashboard visualization and alerting	Real-time stock and reorder notifications

Source: authors' compilation

3. Method

This research adopts an applied, quantitative research design based on the Knowledge Discovery in Databases (KDD) process, adapted to produce a digital sales forecasting model and its associated inventory-control outputs. The overall research stages are literature review and field study, data selection, data preprocessing, data transformation, train-test split, data mining (model development and evaluation), and calculation of safety stock and reorder point, as summarized in Table 5.

Table 5: Research Stages

No.	Stage	Description
1	Literature review & field study	Review prior research; interviews and documentation with Company staff
2	Data selection	Identify relevant data sources: daily sales and national holiday data
3	Data preprocessing	Clean data; remove duplicates and missing values; feature engineering
4	Data transformation	Normalize and convert data into a format suitable for modeling
5	Train-test split	Divide data into training and testing subsets
6	Data mining	Train and evaluate the XGBoost model using RMSE, MAPE, and cross-validation.
7	Inventory-control calculation	Calculate safety stock and reorder point from forecast output

Source: authors' compilation

3.1. Research Setting and Data Source

The research used data from two outlets of a fast-food franchise Company that operates more than 300 outlets in Indonesia; both outlets are anonymized throughout this paper as Outlet A and Outlet B to preserve the confidentiality of the business and its internal operating data. Primary data were obtained through interviews with the Company's operational staff to understand the business problem in depth. In contrast, secondary data consisted of daily sales records for both outlets from January 2023 to June 2024, together with national holiday data collected from publicly available calendars. Supporting data on stock-related customer complaints was gathered from publicly posted customer reviews to corroborate the operational problem described by staff.

As shown later in Figures 1 and 2, daily sales at both outlets fluctuated considerably within the observation period, ranging from roughly 200 to over 600 units per day, with visible spikes consistent with holiday periods and localized demand shocks. This level of volatility is precisely the type of pattern that simple moving-average forecasts tend to underfit, motivating a machine learning approach that can learn from multiple interacting features rather than a single trend line.

3.2. Feature Engineering and Data Preprocessing

The raw daily sales data was first cleaned to remove duplicate records and address missing values. New features were then engineered from the date field so that the model can capture repeating demand patterns, including day-of-week and month indicators and a binary national holiday flag, as weekly sales of fast food in this category tend to show both seasonality (e.g., increased weekend sales) and holiday-related demand shocks. This step is consistent with common practice in applied sales-forecasting research, where date-derived features consistently improve model accuracy relative to a raw time index. (Raizada & Saini, 2021).

3.3. Data Transformation and Train-Test Split

The cleaned and feature-engineered dataset was converted into a format suitable for the XGBoost algorithm, including any encoding required for categorical features such as day-of-week and holiday flags. It was then divided into training and testing subsets, with the training subset used to fit the model and the testing subset reserved to evaluate out-of-sample forecasting performance for the July–December 2024 period, as reported in the Findings and Discussion sections.

3.4. Model Development and Evaluation

The XGBoost regression model was implemented in Python using Google Colab as the computing environment. Hyperparameters such as learning rate, maximum tree depth, and number of boosting rounds were tuned to minimize validation error, as is standard practice for gradient boosting models applied to tabular sales data. Model performance was assessed using five-fold cross-validation, with RMSE and MAPE computed for each fold and averaged to obtain the overall metrics reported in Tables 7 and 8, ensuring that accuracy is not driven by any single unusually easy or difficult period in the data.

3.5. Safety Stock and Reorder Point Calculation

Using the sales forecast results for the July–December 2024 period, the mean and standard deviation of predicted daily sales were computed for each outlet. Safety stock was then calculated in Microsoft Excel following the formula presented in the Literature Review, based on the standard deviation of predicted sales and an assumed service level appropriate for a fast-moving, perishable product category. The reorder point for fried chicken products was subsequently obtained by combining expected demand during the replenishment lead time with the safety stock value. These calculations translate the statistical forecast output into concrete inventory parameters that Company staff can apply directly in daily stock-ordering decisions.

3.6. Validity and Reliability

Internal validity was supported by using historical, transaction-level sales data rather than self-reported estimates, which reduces recall and social-desirability bias that can affect interview-based demand estimates. Reliability was assessed through five-fold cross-validation, so the reported RMSE and MAPE values reflect performance averaged across several data partitions. External validity is more limited, as discussed in the Limitations section, because the model was trained and evaluated using data from only two outlets of a single franchise brand.

3.7. Data Confidentiality and Anonymization

Because the data used in this study includes proprietary daily sales figures belonging to a private franchise business, the research followed a confidentiality protocol to protect commercially sensitive information. The Company's name, its brand identifiers, and the names of the two outlets studied were replaced throughout this paper with the neutral labels "the Company," "Outlet A," and "Outlet B," and any brand marks or personal names visible in supporting screenshots, such as the dashboard mock-up in Figure 3, were masked prior to inclusion. This approach is consistent with standard practice in applied business research involving proprietary operational data and does not affect the validity of the quantitative results, since all reported sales figures, forecast metrics, and inventory-control values are reported exactly as computed from the underlying data.

Table 6 summarizes the software tools used at each stage of the research process, from data preparation through to inventory-control calculation.

Table 6: Software and Tools Used in This Study

Stage	Tool	Purpose
Data preprocessing & feature engineering	Python (pandas)	Cleaning, transformation, and feature construction
Model development	Python (XGBoost library), Google Colab	Training and validating the forecasting model
Model evaluation	Python (scikit-learn)	Computing RMSE, MAPE, and cross-validation folds
Inventory-control calculation	Microsoft Excel	Calculating safety stock and reorder point
Dashboard concept design	UI design tool	Designing the proposed monitoring dashboard mock-up

Source: authors' compilation

4. Findings

Table 7 presents XGBoost's sales-forecasting performance for Outlet A.

Table 7: Outlet A Sales Forecast Evaluation Results

Cross-validation	RMSE	MAPE
Fold 1	56.94	10%
Fold 2	88.08	13.1%
Fold 3	66.01	9.7%
Fold 4	50.13	9%
Fold 5	48.09	9%
Average	61.89	10.2%

Source: authors' analysis

Across the five validation folds, RMSE ranged from approximately 48 to 88 units, with Fold 2 showing the highest error, consistent with a period of unusually high sales volatility visible as the tallest spikes in the blue historical series in Figure 1. The average MAPE of 10.2% places the forecast just inside the “good forecast” range of 10–20%, indicating that the model captures the outlet’s overall demand pattern reliably, even though a small number of high-volatility days remain harder to predict precisely. Forecasts for this outlet were generated for the July–December 2024 period, shown in Figure 1.

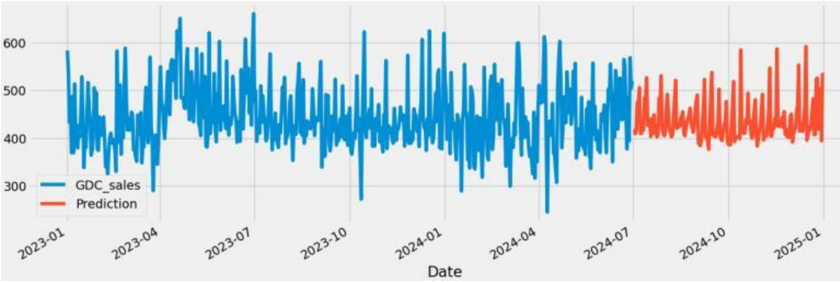

Figure 1: Outlet A Sales Forecasting

Source: authors' analysis

Table 8: Outlet B Sales Forecast Evaluation Results

Cross-validation	RMSE	MAPE
Fold 1	52.26	16.3%
Fold 2	68.52	18.1%
Fold 3	68.03	13.7%
Fold 4	44.34	9.8%
Fold 5	47.82	12.6%
Average	56.2	14.1%

Source: authors' analysis

Table 8 presents XGBoost's performance in forecasting sales for Outlet B. Outlet B shows a higher average MAPE (14.1%) than Outlet A and also more variation across folds, ranging from 9.8% to 18.1%. This is consistent with Figure 2, which shows a pronounced demand spike in April 2024 well above Outlet B's normal sales level; folds containing this period are harder for the model to predict accurately, which may explain the larger error in Folds 1 and 2 than in Folds 4 and 5. Even so, the average MAPE of 14.1% remains within the 10–20% "good forecast" range, indicating that the model still provides a usable planning basis for this outlet despite its higher demand volatility. Forecasts for this outlet were generated for the July–December 2024 period, shown in Figure 2.

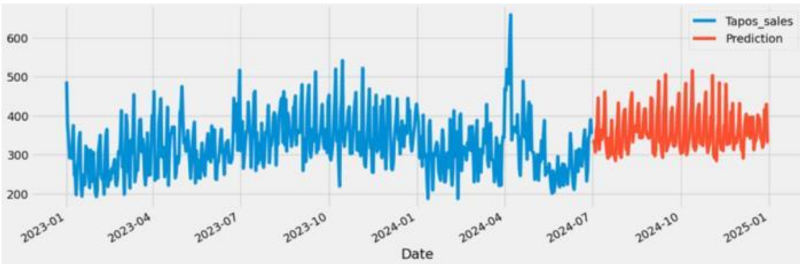

Figure 2: Outlet B Sales Forecasting

Source: authors' analysis

4.1. Comparative Analysis Between Outlets

Comparing the two outlets highlights an important practical point for the Company's digital transformation strategy: forecast accuracy is not uniform across outlets, even under an identical modeling approach. Outlet A, with lower average sales volatility, achieves a lower MAPE (10.2%) than Outlet B (14.1%), whose sales pattern includes a sharper, less predictable demand spike. As the Company scales the forecasting system to additional outlets, it should therefore plan for outlet-specific accuracy levels rather than assuming a single network-wide figure, and treat outlets with higher historical volatility as priority candidates for additional safety stock buffer or closer manual review.

4.2. Safety Stock and Reorder Point

Safety stock and reorder point calculations were based on the sales forecast results described above. The mean and standard deviation of predicted sales for the July–December 2024 period are shown in Table 9.

Table 9: Mean and Standard Deviation of Sales Predictions

Outlet	Average	Standard Deviation
Outlet A	435.35	42.15
Outlet B	359.67	51.89

Source: authors' analysis

Based on these values, Table 10 shows the resulting safety stock and reorder point for each outlet.

Table 10: Safety Stock and Reorder Point

Outlet	Safety Stock	Reorder Point
Outlet A	98	968
Outlet B	120	840

Source: authors' analysis

Outlet B has both a lower average predicted demand (359.67 units) than Outlet A (435.35 units) and a higher standard deviation (51.89 versus 42.15), which explains why Outlet B requires a proportionally larger safety stock (120 units, or about 33% of its average daily demand) than Outlet A (98 units, or about 23%). Outlet B's higher demand variability, evident in both its MAPE and its standard deviation, translates directly into a higher relative buffer requirement, precisely the kind of outlet-specific insight that a centralized, non-forecast-based inventory policy would be unlikely to capture.

4.3. Digital Business Recommendations

Building on these results, five digital-business-oriented recommendations are proposed for the Company's Outlet A and Outlet B, and more broadly for its outlet network—implementation of the sales forecasting system in production. The forecasting model developed in this study should move from a research prototype into an operational tool that is retrained periodically as new sales data becomes available, so that its forecasts remain accurate as demand patterns evolve. Development of a visualization dashboard. The dashboard, illustrated conceptually in Figure 3, would enable outlet and area managers to monitor stock availability, sales, forecasts, and reorder alerts in real time, removing the need to track spreadsheets manually and allowing faster reaction to low-stock alerts and the development of promotional strategies for periods of expected demand increase. Because the forecasting model can identify future high-demand periods such as national holidays, the Company can schedule promotions and staffing in advance rather than reacting to demand after it occurs. Monitoring client feedback on an ongoing basis. Regular review of client feedback, including stock-out-related online reviews, should serve as a qualitative check on the quantitative forecasting and inventory-control system, helping the Company identify service gaps that inventory metrics alone may not reveal.

Staff training and change management. Because the system changes how outlet staff plan daily production and ordering, the Company should invest in structured training and a change-management process so that staff trust and correctly interpret the dashboard's forecasts and alerts, rather than reverting to informal manual estimates.

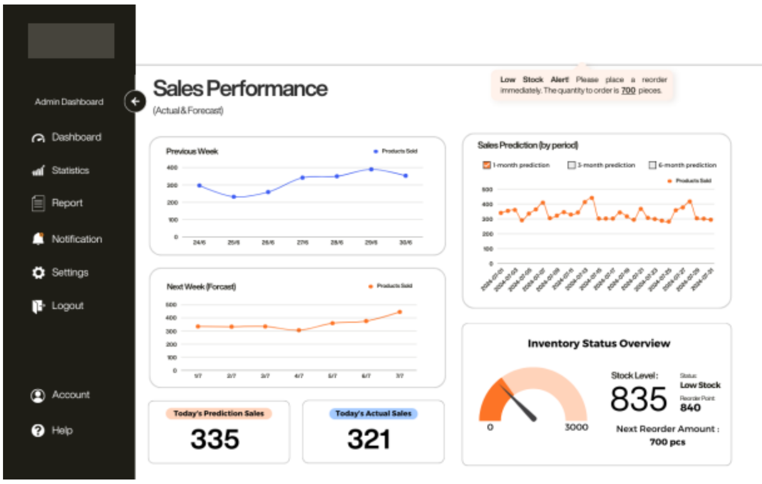

Figure 3: Conceptual Design of the Sales and Inventory Monitoring Dashboard

Source: authors' compilation

Table 11 consolidates these five recommendations with a suggested implementation priority and responsible party, to support translating them from a research proposal into an actionable rollout plan.

Table 11: Implementation Summary of Digital Business Recommendations

Recommendation	Suggested Priority	Primary Responsible Party
Production deployment of forecasting system	High	Company IT/data team
Real-time monitoring dashboard	High	Company IT team with outlet manager input
Demand-aligned promotional planning	Medium	Marketing/promotions team
Customer feedback monitoring	Medium	Customer service/quality assurance team
Staff training and change management	High	Operations / human resources team

Source: authors' compilation

5. Discussion

The findings show that the XGBoost-based forecasting model produced usable outlet-level predictions for both outlets, with average MAPE values of 10.2% for Outlet A and 14.1% for Outlet B. Both values fall within the commonly accepted good-forecast range, indicating that the model can support operational planning even in a fast-food franchise environment characterized by daily demand fluctuation and perishable inventory constraints.

These results extend prior forecasting and inventory-management studies by demonstrating an integrated forecasting-to-inventory workflow in a franchise-based, perishable-product context. The findings suggest that machine learning can provide value not only through prediction accuracy, but also through its ability to translate historical sales data into actionable inventory parameters and operational recommendations.

5.1. Theoretical Contribution

From a theoretical perspective, this study contributes to the sales-forecasting and inventory-management literature by demonstrating how a machine learning forecast can be directly linked to safety stock and reorder point calculation within a single operational workflow, rather than treating forecasting and inventory control as separate research problems, as much of the prior literature summarized in Table 2 tends to do (Heaviside et al., 2020; Nissa et al., 2023). The study thereby offers a template adaptable to other perishable-product or franchise settings.

5.2. Practical Contribution and Forecast Accuracy in Context

From a practical standpoint, the study contributes a concrete digital transformation pathway for franchise businesses: forecasting output feeds directly into inventory parameters and into a proposed monitoring dashboard, giving managers an integrated, data-driven view of demand and stock position instead of relying on manual estimation. For a franchise such as the Company, where fixed food-safety time limits leave little room for error in production planning, this integration offers a direct route to reducing both overstock and understock incidents. At the same time, outlet-level comparisons show where additional buffer stock or managerial attention is most needed.

The accuracy achieved here (MAPE of 10.2% and 14.1%) compares favorably with several previously reported forecasting accuracy levels in adjacent retail contexts. (Huo, 2021; Panarese et al., 2022). This indicates that XGBoost is a reasonable default choice for franchise businesses starting their forecasting digitalization journey, especially where historical data exists, but resources for extensive model experimentation are limited.

5.3. Cost-Efficiency and Industry Implications

From a cost-efficiency perspective, even a partial reduction in the Company's reported baseline overstock rate of approximately 30 products per day would be operationally meaningful at scale, given its network of more than 300 outlets. Because overstocked fried chicken must be withdrawn under the FIFO food-safety policy rather than discounted, every overstocked unit represents a full loss of the input cost, so outlet-level accuracy improvements scale multiplicatively into network-wide waste reduction if the model is extended beyond the two outlets studied. A similar logic applies to understock: each stockout is both a lost sale and a source of negative online reviews that reach customers who were never present at the outlet.

Beyond the Company's own operations, the findings carry implications for how franchise associations, industry groups, and policymakers might support digital transformation among smaller franchise businesses. Many operators, particularly smaller regional chains, already possess the transaction-level point-of-sale data used here but lack the analytical capacity to convert it into such a workflow. Industry associations and business incubators could disseminate accessible, low-cost forecasting templates so that smaller operators are not excluded from the efficiency gains available to larger, better-resourced chains.

5.4. Managerial Implications

From a managerial-practice standpoint, three implications follow from the outlet-level comparison. First, outlet-specific monitoring matters: because Outlet B's volatility raised both its MAPE and its safety stock requirement, a single network-wide buffer policy would over-supply stable outlets such as Outlet A while under-supplying volatile ones such as Outlet B. Second, forecast-driven promotional planning is most valuable at the outlets and periods with the largest predicted demand increases, allowing marketing spend to be matched by adequate stock. Third, change management is integral to any forecasting rollout, since even an accurate model has limited value if staff continue to rely on informal estimates.

Table 12 summarizes these implications by operational indicator, framed qualitatively rather than as numerical projections, since this study did not conduct a controlled before-and-after deployment.

Table 12: Qualitative Expected Impact of Digital Forecasting Adoption on Key Operational Indicators

Operational Indicator	Expected Direction of Change	Primary Mechanism
Overstock incidents (units withdrawn under FIFO policy)	Decrease	More accurate daily production quantities from forecast output
Understock/stockout incidents	Decrease	Reorder point calculated from forecast mean and variability
Negative customer reviews related to stock issues	Decrease	Fewer stockouts reduce a key source of customer dissatisfaction
Manual planning time spent by outlet supervisors	Decrease	Dashboard automates calculations previously done informally
Consistency of inventory decisions across outlets	Increase	Shared, standardized forecasting and calculation workflow
Responsiveness to demand spikes (e.g., holidays)	Increase	Holiday flag feature allows the model to anticipate spikes in advance

Source: authors' compilation

6. Conclusion

This study developed an XGBoost-based digital sales forecasting model for two outlets of an Indonesian fast food franchise, referred to as Outlet A and Outlet B to preserve business confidentiality, producing forecasts categorized as good (MAPE of 10.2% for Outlet A and 14.1% for Outlet B). Using these forecasts, the study calculated safety stock and reorder point values for each outlet, showing that the outlet with higher demand volatility (Outlet B) required a proportionally larger safety stock buffer than the more stable outlet (Outlet A), a distinction that a non-forecast-based inventory policy would likely miss. These technical results were translated into five practical digital-business recommendations: production deployment of the forecasting model, a real-time monitoring dashboard, demand-aligned promotional planning, systematic customer-feedback monitoring, and staff training to support adoption.

Taken together, these findings show that combining machine learning-based forecasting with classical inventory-control calculations, within a single integrated digital workflow, can provide franchise fast food businesses with an effective, technology-enabled solution to the overstock and understock problems that frequently affect their operations. Beyond the two outlets studied, this research offers a practical, replicable reference for other franchise and small culinary businesses seeking digital transformation in their inventory and production planning.

Returning to the four research questions posed at the outset, the XGBoost-based model forecast daily sales with a MAPE of 10.2% at Outlet A and 14.1% at Outlet B, both within the "good forecast" range (RQ1); applying standard inventory-control formulas to this output yielded a safety stock of 98 units and a reorder point of 968 units for Outlet A, and 120 and 840 units for Outlet B (RQ2); accuracy and inventory parameters differed between the outlets, with the more volatile outlet showing both a higher MAPE and a larger buffer requirement, underscoring the value of outlet-specific policies (RQ3); and five digital-business recommendations were developed to translate these technical results into practice (RQ4).

7. Limitations and Future Study

This study has several limitations. First, the model was built using data from only two outlets over roughly a year and a half, which may limit generalizability to the Company's other outlets or to franchises with different demand patterns, store formats, or regional characteristics. Second, the forecasting model relies primarily on historical sales and holiday data and does not incorporate other potentially influential variables such as pricing changes, local promotions, competitor actions, or weather conditions. Third, the study applies a single algorithm, XGBoost, without a systematic comparative benchmark against alternatives such as ARIMA, exponential smoothing, or neural network-based approaches. Hence, its relative advantage for this specific business case remains to be confirmed directly on this dataset. Fourth, the safety stock and reorder point calculations rely on service-level and lead-time assumptions that were appropriate for this case but were not varied systematically. Hence, their sensitivity was not formally tested.

Future research could extend this model to a larger number of outlets to test its scalability and to examine whether outlet clustering, based on demand-pattern similarity, could allow a smaller number of shared models rather than one model per outlet. Future work could also incorporate additional demand drivers such as promotional calendars, local events, or weather data, and empirically compare XGBoost against alternative machine learning or statistical forecasting methods on the same dataset. In addition, future research could move beyond the conceptual dashboard proposed here toward a deployed, real-time monitoring system, and evaluate its actual effect on overstock and understock rates, staff decision-making, and customer satisfaction once implemented in production, ideally using a pre-post or controlled comparison design across a subset of outlets.

Author Contributions

Conceptualization: Meinarini Catur Utami and Mim Hanifah Permana; Methodology: Meinarini Catur Utami and Elvi Fetrina; Data Collection: Mim Hanifah Permana; Formal Analysis: Mim Hanifah Permana and Meinarini Catur Utami; Writing – Original Draft Preparation: Mim Hanifah Permana; Writing – Review and Editing: All authors; Supervision: Meinarini Catur Utami and Elvi Fetrina. All authors have read and approved the final version of the manuscript.

Declarations

Conflict of Interest: The authors declare that there are no conflicts of interest regarding the publication of this article.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability: The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Ethical Approval: This study was conducted in accordance with ethical research principles involving human participants. Participation was voluntary, respondents were informed about the purpose of the research, and informed consent was obtained before data collection. The anonymity and confidentiality of all respondents were maintained throughout the research process.

Use of AI Tools: The authors declare that artificial intelligence (AI) tools were used solely for language enhancement, grammar checking, and manuscript editing assistance. All conceptual development, data analysis, interpretation of results, and final academic content remain the sole responsibility of the authors.

References

Aghnani, K., & Perdana, T. (2019). Memahami Manajemen Persediaan pada Perusahaan Agroindustri dengan Menggunakan Alat Berpikir Sistematis. *Jurnal Ilmiah Mahasiswa Agroinfo Galuh*, *6*(1), 259–272. https://doi.org/10.25157/jimag.v6i1.1950

Allied Market Research, R. (2020). *Fast Food Market Size, Share | Growth Opportunities – 2027*. Allied Market Research. https://www.alliedmarketresearch.com/fast-food-market

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16*, 785–794. https://doi.org/10.1145/2939672.2939785

Coherent Market Insight, C. M. I. P. (2026, March 26). *Fast Food Market Size, Trends & YoY Growth Rate, 2026-2033*. Coherent Market Insights. https://www.coherentmarketinsights.com/market-insight/fast-food-market-5896

Dairu, X., & Shilong, Z. (2021). Machine Learning Model for Sales Forecasting by Using XGBoost. *2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE)*, 480–483. https://doi.org/10.1109/ICCECE51280.2021.9342304

Ekankumo, B. (2023). Marketing Mix Elements and Brand Loyalty in the Nigerian Telecommunication Industry. *British Journal of Management and Marketing Studies*, *6*(1), 84–96. https://doi.org/10.52589/BJMMS-YXXASLJM

Hasan, M. R., 1 1 M. B. A. (2024). *Addressing Seasonality and Trend Detection in Predictive Sales Forecasting: A Machine Learning Perspective*. 100–109. https://doi.org/10.32996/jbms.2024.6.2.10

Heaviside, M., Mulyawan, B., & Sutrisno, T. (2020). Determination of minimum stock on system retail using forecast, economic order quantity, and reorder point methods. *IOP Conference Series: Materials Science and Engineering*, *1007*(1), 012180. https://doi.org/10.1088/1757-899X/1007/1/012180

Huo, Z. (2021). Sales Prediction based on Machine Learning. *2021 2nd International Conference on E-Commerce and Internet Technology (ECIT)*, 410–415. https://doi.org/10.1109/ECIT52743.2021.00093

Husein, A. M., Lubis, F. R., & Harahap, M. K. (2021). Analisis Prediktif untuk Keputusan Bisnis: Peramalan Penjualan. *Data Sciences Indonesia (DSI)*, *1*(1), 32–40. https://doi.org/10.47709/dsi.v1i1.1196

Megawati, S., Herdiansyah, H., Machmud, A., Antriyandarti, E., & Sudirman, S. (2024). Integrating circular economy, digital economy, and social protection policies to drive green business innovation: Insights from Indonesia's culinary SMEs. *Problems and Perspectives in Management*, *22*(4), 368–381. https://doi.org/10.21511/ppm.22(4).2024.28

Nabrawi, E., & Alanazi, A. (2023). Fraud Detection in Healthcare Insurance Claims Using Machine Learning. *Risks*, *11*(9), 160. https://doi.org/10.3390/risks11090160

Nissa, D. A., Supian, S., & Nahar, J. (2023). Inventory Control for MSME Products Using the Q Model with Lost Sales Condition Based on Product Sales Forecasting. *International Journal of Quantitative Research and Modeling*, *4*(1), 20–29. https://doi.org/10.46336/ijqrm.v4i1.417

Panarese, A., Settanni, G., Vitti, V., & Galiano, A. (2022). Developing and Preliminary Testing of a Machine Learning-Based Platform for Sales Forecasting Using a Gradient Boosting Approach. *Applied Sciences*, *12*(21), 11054. https://doi.org/10.3390/app122111054

Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M. Z., Barrow, D. K., Ben Taieb, S., Bergmeir, C., Bessa, R. J., Bijak, J., Boylan, J. E., Browell, J., Carnevale, C., Castle, J. L., Cirillo, P., Clements, M. P., Cordeiro, C., Cyrino Oliveira, F. L., De Baets, S., Dokumentov, A., … Ziel, F. (2022). Forecasting: Theory and practice. *International Journal of Forecasting*, *38*(3), 705–871. https://doi.org/10.1016/j.ijforecast.2021.11.001

Putra, S. A. (2019). Analisis Peramalan Penjualan dan Promosi Penjualan terhadap Peningkatan Volume Penjualan pada PT. Cakra Anugerah Arta Alumindo Medan. *Jurnal Manajemen Dan Akuntansi Medan*, *1*(3), 142–160. https://doi.org/10.47709/jumansi.v1i3.2097

Raizada, S., & Saini, J. R. (2021). Comparative Analysis of Supervised Machine Learning Techniques for Sales Forecasting. *International Journal of Advanced Computer Science and Applications (IJACSA)*, *12*(11). https://doi.org/10.14569/IJACSA.2021.0121112

Turgut, Y., & Erdem, M. (2022). Forecasting of Retail Produce Sales Based on XGBoost Algorithm. In F. Calisir (Ed.), *Industrial Engineering in the Internet-of-Things World* (pp. 27–43). Springer International Publishing. https://doi.org/10.1007/978-3-030-76724-2_3

Wanti, L. P., Maharrani, R. H., Prasetya, N. W. A., Tripustikasari, E., & Ikhtiagung, G. N. (2020). Optimization of the economic order quantity method for a support system reorder point stock. *International Journal of Electrical and Computer Engineering (IJECE)*, *10*(5), 4992–5000. https://ijece.iaescore.com/index.php/IJECE/article/view/22000