Multidisciplinary Reviews

REVIEW ARTICLE
Published Online: August 26, 2025
https://doi.org/10.31893/multirev.2026127

Machine learning applied to sales prediction modeling: A systematic literature review

Singgih Saptadi[a] | **Wiwik Budiawan**[a] | **Rifky Nafi Fadhila**[a] | **Rangga Kamajaya Sumbodo**[a] | **Dhimas Wachid Nur Saputra**[a]

[a]Department of Industrial Engineering, College of Industrial Engineering and Management, Diponegoro University, Indonesia.

Abstract Sales prediction plays a critical role across industries by enabling organizations to make informed decisions in inventory control, marketing, and strategic planning. This study aims to systematically review the application of machine learning (ML) techniques in sales forecasting, offering a synthesized understanding of prevailing algorithms, evaluation metrics, sector-specific challenges, and the distinction between theoretical and practical research. Adopting the PRISMA-P methodology, a structured literature search was conducted on the Scopus database, covering studies published from 2014 to 2024. From 206 initial publications, 66 were selected after applying inclusion criteria and quality assessment procedures. Findings show that ensemble learning techniques, particularly XGBoost and Random Forest, are most frequently used and consistently outperform traditional regression models in predictive accuracy. Common evaluation metrics include Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R-squared (R^2), with RMSE being the dominant choice. Sector-specific insights reveal distinct forecasting challenges: the retail and e-commerce sectors grapple with demand volatility and rapidly shifting consumer behavior, the manufacturing industry contends with supply chain complexities and economic disruptions, while the energy sector deals with high-frequency fluctuations and data granularity issues. Moreover, this review identifies a clear divide between model-driven studies focused on algorithmic optimization and problem-based research that integrates contextual variables such as consumer sentiment, macroeconomic trends, and operational constraints to enhance real-world applicability. The review highlights that hybrid and adaptive approaches—especially those incorporating external data and collaborative industry inputs—offer promising pathways toward more actionable and scalable forecasting solutions. By bridging the gap between technical performance and practical deployment, this systematic literature review underscores the importance of aligning ML methodologies with real-world business contexts to advance the effectiveness of sales prediction systems.

Keywords: sales prediction, forecasting, machine learning, PRISMA-P, systematic literature review

1. Introduction

Sales prediction has emerged as a crucial application of machine learning (ML) in various industries, helping businesses optimize inventory management, marketing strategies, and overall decision-making. The ability to forecast future sales accurately can provide companies with a competitive advantage by reducing costs, minimizing stock shortages or overstock situations, and enhancing customer satisfaction. Traditional statistical forecasting techniques, such as moving averages and exponential smoothing, have been widely used for sales prediction. However, these approaches often struggle to capture complex, nonlinear patterns in consumer behavior, seasonal trends, and external market factors (Ali et al., 2023; Branda et al., 2020; Fourkiotis & Tsadiras, 2024). As a result, machine learning techniques have gained popularity for their ability to analyze large datasets and uncover hidden patterns, leading to more accurate and adaptive forecasting models (Han & Li, 2024; Wisesa et al., 2020a).

The application of ML in sales prediction has been explored across multiple sectors, including retail (Fotache et al., 2021; Li et al., 2023; Rathore et al., 2024), food and beverage (Wang, 2020), pharmaceuticals (Samara & Stanich, 2023), and sports event ticketing (Chalapathy et al., 2022). Each sector presents unique challenges, such as fluctuating demand, seasonality, and external influences like economic conditions and competitor actions. Consequently, different ML models have been employed to address these sector-specific complexities. Regression models, time series forecasting techniques, and ensemble learning methods are among the most commonly used approaches (Akanksha et al., 2022; Dutta et al., 2022; Singh et al., 2020). While some studies demonstrate the effectiveness of traditional models like ARIMA and Holt-Winters for capturing trends (Samara & Stanich, 2023), others highlight the superior performance of advanced methods such as XGBoost, LightGBM, and deep learning-based models (Li et al., 2023).

Given the diversity of ML models available for sales forecasting, it is essential to assess their effectiveness based on standard performance metrics. Evaluation measures such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE),

 Received: May 18, 2025 | Accepted: July 6, 2025

Mean Absolute Percentage Error (MAPE), and R-squared (R^2) are widely used to quantify prediction accuracy (Kamble et al., 2023; Patil et al., 2023; Rathore et al., 2024). These metrics help researchers and practitioners determine the reliability of different models under various conditions. However, the choice of an appropriate evaluation metric depends on the nature of the dataset and the specific forecasting goals of the study (Akanksha et al., 2022; Kamble et al., 2023).

A few studies emphasize the importance of integrating external factors such as macroeconomic indicators, consumer sentiment, and competitive market dynamics into sales prediction models to enhance their robustness and applicability (Achumie et al., 2022; Ching et al., 2018; Lin et al., 2024). Incorporating external data sources, such as social media trends, online reviews, and web search activity, has shown potential in improving forecasting accuracy by capturing real-time shifts in consumer preferences and market conditions (Ibrahim & Wang, 2019; Jun et al., 2014). Additionally, hybrid models that combine traditional statistical techniques with machine learning approaches have gained traction for their ability to balance interpretability and predictive power (Akusta, 2024; Dritsas & Trigka, 2025). However, challenges such as data privacy concerns, computational complexity, and the need for extensive feature engineering remain key obstacles to broader adoption (Elhanashi et al., 2023; Rane et al., 2024; Rozony et al., 2024). Addressing these challenges is essential for advancing ML-driven sales prediction research and ensuring its practical relevance in dynamic business environments.

Despite advancements in ML-based sales prediction, several challenges persist. Data availability, quality, and feature selection significantly impact model performance (Li et al., 2023). Additionally, the scalability and generalizability of ML models remain critical concerns, as models trained on historical sales data may fail to adapt to sudden market changes, economic disruptions, or new consumer behaviors (Chalapathy et al., 2022). Furthermore, research in sales forecasting often focuses on technical improvements in prediction accuracy rather than addressing real-world business challenges (Patil et al., 2023). Distinguishing between studies that offer practical business insights and those that merely enhance predictive performance is essential for bridging the gap between academic research and industry applications (Rathore et al., 2024).

This Systematic Literature Review (SLR) aims to synthesize existing research on ML-based sales prediction by addressing four key research questions:

RQ1: What machine learning methods are used for sales prediction?
RQ2: What performance metrics evaluate machine learning models in sales prediction?
RQ3: Which sectors face challenges related to sales forecasting?
RQ4: How can we distinguish research that addresses real problems from research that focuses only on sales prediction?

2. Materials and Methods

This study employs the PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols) method to systematically identify, screen, and select relevant publications for review. PRISMA-P is a standardized approach designed to enhance transparency, consistency, and quality in systematic reviews and meta-analyses. Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols (PRISMA-P) consist of several key stages: identification, screening, eligibility, and inclusion. As shown in Figure 1, each step ensures that only the most relevant studies contribute to answering the research questions.

The identification stage is the first step in the PRISMA-P methodology, where relevant literature is systematically searched and collected. This study focused on retrieving academic publications from the Scopus database due to its extensive coverage of high-quality, peer-reviewed research across various disciplines. This paper use search string (("sale prediction" OR "sales predictions") AND ("machine learning" OR "machine learning technique")) to capture studies discussing sales prediction using machine learning techniques, ensuring that a broad yet relevant range of research papers was included in the initial collection. The search was limited to a publication range from 2014 to 2024, allowing the inclusion of recent advancements in machine learning techniques applied to sales forecasting while maintaining a long enough timeframe to analyze research trends over a decade. A total of 206 publications were identified in Scopus. These publications included various types of academic contributions, such as conference papers, journal articles, reviews, book chapters, and errata. This initial collection represents the broadest dataset before further refinement in the subsequent screening stage.

The screening stage involved filtering the identified publications based on predefined inclusion and exclusion criteria to ensure that only the most relevant and high-quality studies were selected for further analysis. The inclusion and exclusion criteria were designed to focus on studies presented in Table 1.

After applying the screening criteria, the 206 initially collected papers were classified into 128 conference papers, 57 journal articles, 9 conference reviews, 7 book chapters, 3 review articles, and 2 errata. Papers that did not meet the inclusion criteria were removed. Specifically, 9 conference reviews, 7 book chapters, 3 review articles, and 2 errata were excluded from further analysis. This filtering resulted in a total of 104 relevant papers, consisting of 73 conference papers and 31 journal articles, which were then analyzed in more depth in the eligibility stage.

The next step is eligibility. At this stage, the 104 publications that passed the screening process were thoroughly analyzed to assess their relevance and quality. Papers were evaluated based on their contribution to answering the research questions, particularly in terms of machine learning techniques, performance metrics, and real-world challenges in sales forecasting. Following this analysis, 66 papers (51 conference papers and 15 journal articles) were selected for inclusion in the final review. These studies form the foundation for the systematic literature review, providing insights into the various machine learning methodologies, evaluation metrics, industry challenges, and distinctions between theoretical and practical research approaches.

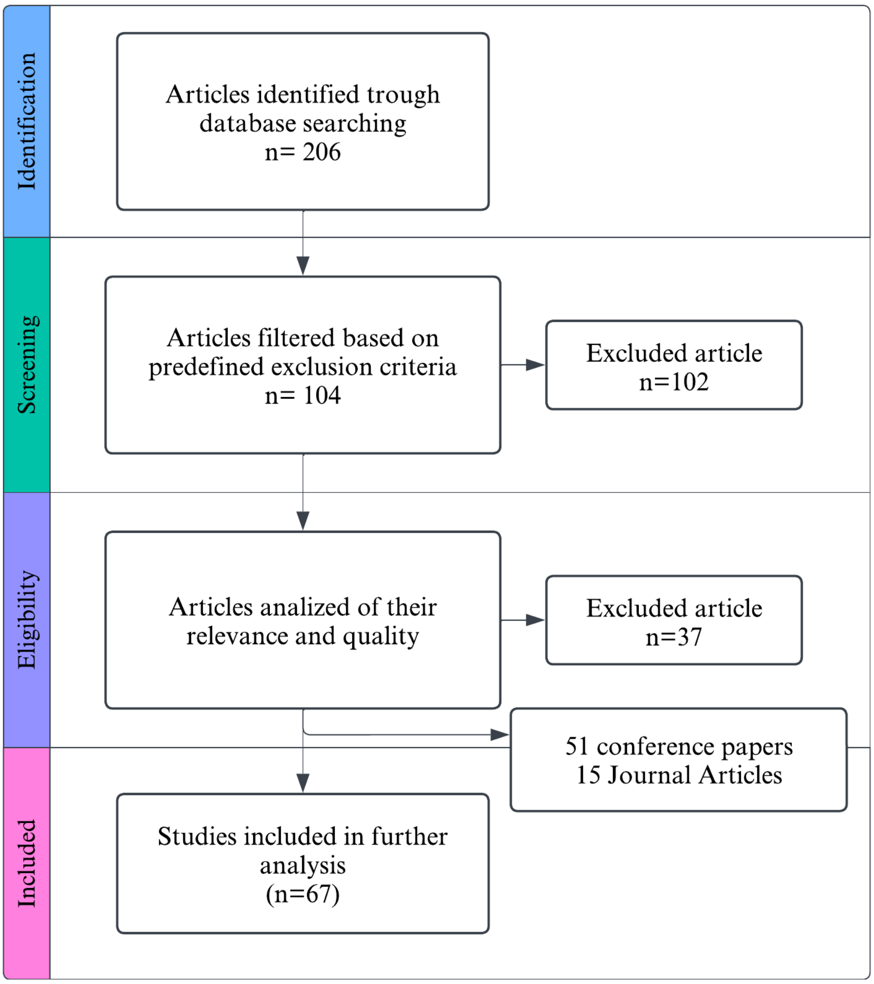

Figure 1 Prisma-P methodology steps followed in this study.

Table 1 Inclusion and exclusion criteria.

Inclusion Criteria
Research that presents machine learning algorithms and performance metrics
Research that focuses on solving problems related to sales prediction
Presentation of research using English
Research that can be fully accessed
Research that is included in scientific papers
Exclusion Criteria
Research that does not present machine learning algorithms and performance metrics
Research that focuses on developing models related to sales prediction
Research presentation that does not use English
Research that cannot be fully accessed and only displays the abstract.
Research that is not included in scientific papers (such as conference reviews, book chapters, reviews, errata, etc.)

3. Results and Discussion

3.1. Publication distribution over the years

Figure 2 shows the number of articles published between 2014 and 2024 (based on the extraction criteria of this study) with a trend line. The results indicate a growing interest in the application of machine learning for sales prediction. Notably, no publications were found between 2014 and 2016, suggesting that research in this area was either limited or not widely documented during that period. The first recorded publication appeared in 2017, with only one paper, followed by another single publication in 2018, reflecting minimal research activity in the early years. However, from 2019 onward, there is a clear upward trend, with the number of publications increasing to four in 2019 and eight in 2020. Although there was a slight decrease in 2021 (7 papers), the numbers remained stable in 2022 (8 papers). A significant surge occurred in 2023, with 21 publications, the highest recorded in this dataset. In 2024, the number of publications remains high at 17, indicating continued research interest. This growing trend may be attributed to advancements in machine learning techniques, increased availability of sales data, and the rising demand for accurate predictive models in various industries.

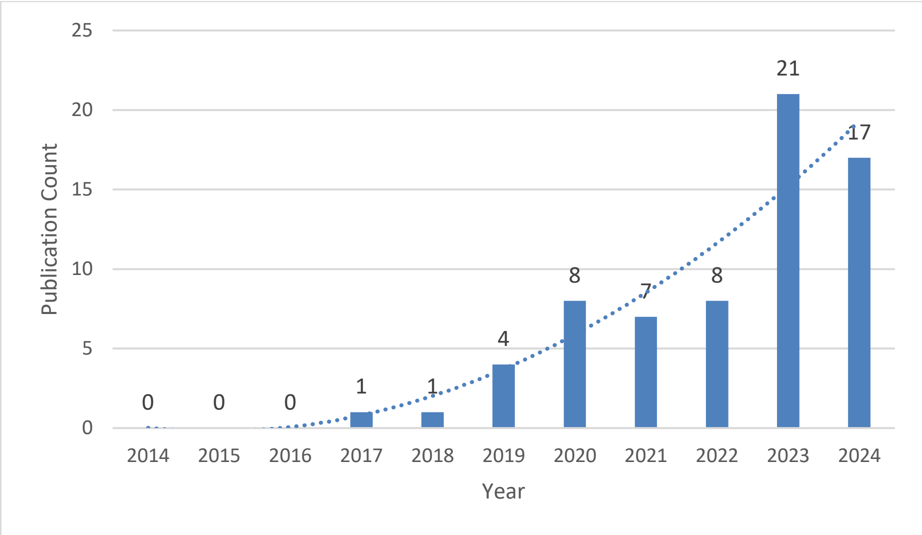

Figure 2 Number of paper per year. *Source:* Scopus Database (2025).

3.2. Machine learning methods used for sales prediction

Machine learning (ML) techniques have become essential tools for sales prediction, offering improved accuracy and adaptability compared to traditional statistical models. Various ML models have been applied in sales forecasting, ranging from simple regression-based methods to more sophisticated ensemble learning techniques and deep learning models as shown in Figure 3 with legend supported in Table 2. Diverse range of approaches categorized into regression models, ensemble methods, deep learning, clustering, and classification techniques. The selection of these methods depends on the nature of the dataset, the complexity of the relationships among variables, and the specific business context of the study.

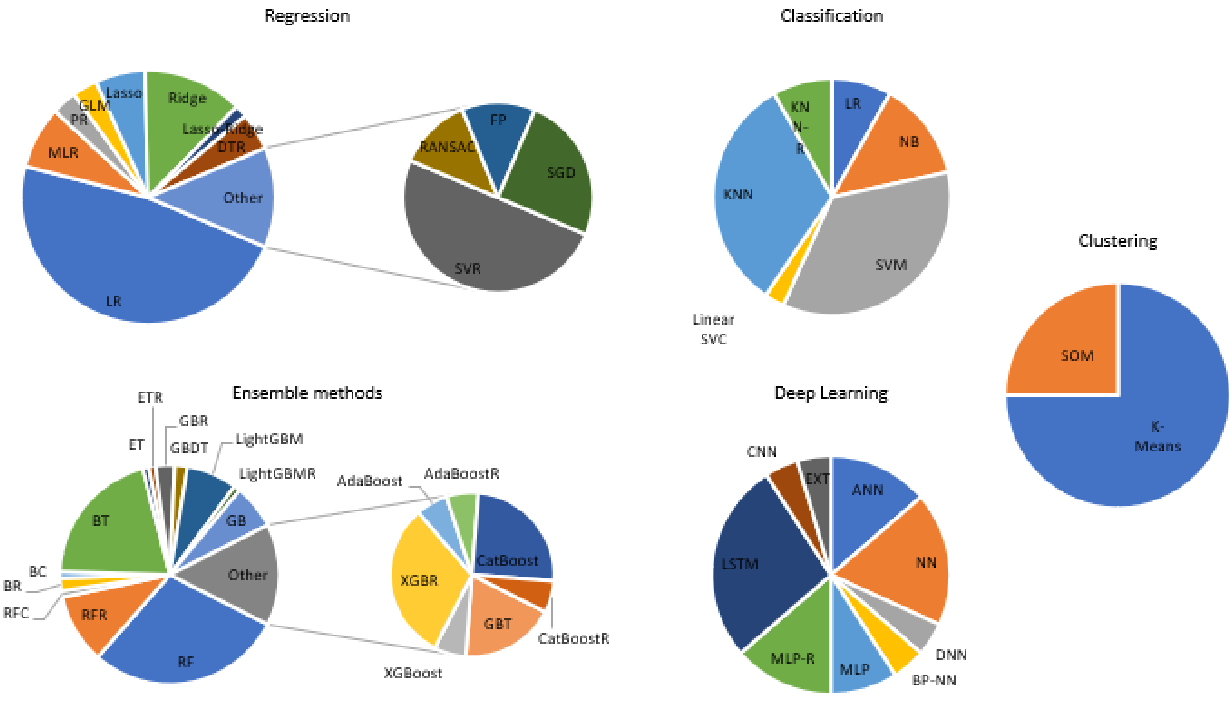

Figure 3 ML methods used for sales predictions. *Source:* Scopus Database (2025).

Table 2 Legend of machine learning methods.

LR	:	Linear Regression	GBT	:	Gradient Boosted Tree
MLR	:	Multiple Linear Regression	XGBoost	:	Extreme Gradient Boosting
PR	:	Polynomial Regression	XGBR	:	Extreme Gradient Boosting Regression
GLM	:	Generalized Linear Model	AdaBoost	:	Adaptive Boosting
Lasso	:	Lasso Regression	AdaBoostR	:	AdaBoost Regressor
Ridge	:	Ridge Regression	CatBoost	:	Categorical Boosting
Lasso-Ridge	:	Lasso Ridge Regression	CatBoostR	:	Categorical Boosting Regression
DTR	:	Decision Tree Regressor	DRF	:	Distributed Random Forest
SVR	:	Support Vector Regression	ANN	:	Artificial Neural Network
RANSAC	:	RANSAC Regression	NN	:	Neural Network
FP	:	Facebook Prophet	DNN	:	Deep Neural Network
SGD	:	Stochastic Gradient Descent	BP-NN	:	Back Propagation Neural Network
RF	:	Random Forest	MLP	:	Multilayer Perceptron
RFR	:	Random Forest Regressor	MLP-R	:	Multilayer Perceptron Regression
RFC	:	Random Forest Classification	LSTM	:	Long Short-Term Memory
BR	:	Bagging Regressor	CNN	:	Convolutional Neural Network
BC	:	Bagging Classifier	EXT	:	Extreme Learning Machine
BT	:	Bagged Tree	K-Means	:	K-Means Clustering
ET	:	Extra Tree	SOM	:	Self-Organizing Map Clustering
ETR	:	Extra Tree Regressor	Ward	:	Ward Clustering
GBR	:	Gradient Boosting Regressor	NB	:	Naïve Bayes
GBDT	:	Gradient Boosting Decision Tree	SVM	:	Support Vector Machine
LightGBM	:	Light Gradient Boosting Machine	Linear SVC	:	Linear Support Vector Machine
LightGBMR	:	Light Gradient Boosting Regression	KNN	:	K-Nearest Neighbors
GB	:	Gradient Boosting	KNN-R	:	K-Nearest Neighbors Regressor

Regression models are among the most commonly used methods in sales prediction. Linear regression is often employed due to its simplicity and interpretability, which serves as a baseline model for predicting sales based on historical data and key influencing factors (Fourkiotis & Tsadiras, 2024; Madhuvanthi et al., 2019; Rathore et al., 2024; Xavier et al., 2023). It has been used in studies analyzing retail sales, such as Black Friday sales forecasting, where its ability to capture general trends in sales data proved beneficial (Rathore et al., 2024). However, linear regression has limitations in capturing nonlinear relationships, which has led researchers to explore alternative approaches. Decision tree regression offers an improvement by segmenting the data into hierarchical rules, which can capture complex interactions among features (Rathore et al., 2024). Meanwhile, K-nearest neighbors (KNN) regression has also been explored to capture localized trends in sales patterns (Patil et al., 2023).

Tree-based ensemble methods such as Random Forest and XGBoost have demonstrated superior performance in many sales prediction studies. Random Forest, an ensemble of multiple decision trees, has been used in forecasting retail store sales and determining inventory stock levels, showing robustness against overfitting and delivering improved accuracy compared to single regression models (Akanksha et al., 2022). XGBoost, a gradient-boosting algorithm, has consistently outperformed other models due to its ability to refine predictions and reduce error rates iteratively. Studies on e-commerce and retail industries highlight XGBoost's effectiveness in handling large, high-dimensional sales datasets (Chalapathy et al., 2022; Patil et al., 2023). Additionally, ensemble techniques that combine multiple models, such as stacking and voting ensembles, have been explored to enhance predictive accuracy further (Kamble et al., 2023).

Several studies have also explored the role of clustering techniques like K-Means Clustering in sales forecasting, often in combination with regression models (Calitz et al., 2022; Ratre & Jayaraj, 2023). Clustering is particularly useful in segmenting customers based on purchasing behavior, allowing businesses to tailor their sales predictions to different customer groups. For example, research in e-commerce sales forecasting used K-means clustering to group customers based on purchase frequency and spending patterns before applying regression models to predict sales trends for each group (Chalapathy et al., 2022).

Deep learning methods, particularly artificial neural networks (ANNs) and recurrent neural networks (RNNs) have been employed in some sales forecasting studies. These models are well-suited for identifying complex patterns in large-scale, time-dependent sales data. In particular, long short-term memory (LSTM) networks, a type of RNN, have been applied to time-series sales forecasting, leveraging their ability to capture sequential dependencies in historical sales data (Kamble et al., 2023). However, deep learning approaches require significant computational resources and extensive datasets, making them less practical for smaller businesses with limited data availability (Kamble et al., 2023).

The datasets used in sales prediction research vary widely depending on the industry and the scope of the study. Retail and e-commerce sales data are among the most commonly used, with many studies relying on publicly available datasets from Kaggle. The Black Friday sales dataset, which includes customer transactions, product categories, and purchase history, has been extensively used to test ML models, particularly regression-based and ensemble learning methods (Akanksha et al., 2022; Niu, 2020; Patil et al., 2023). Other studies utilize multi-year retail sales datasets, incorporating factors such as store location, seasonal variations, and promotions to improve forecasting accuracy (Kamble et al., 2023). In addition, e-commerce company sales data containing customer segmentation and purchase behavior insights have been analyzed using clustering techniques and decision tree-based models (Mallik et al., 2022).

Beyond retail, other industries have also been explored in sales forecasting research. For example, a study on sports ticket sales forecasting used historical match attendance records, weather conditions, and team performance data to predict ticket demand. This research found that Random Forest and XGBoost performed well, demonstrating the adaptability of these models beyond traditional retail applications (Fotache et al., 2021). Similarly, some studies examine grocery store sales, focusing on daily and weekly sales fluctuations to optimize inventory and demand planning (Wen et al., 2024).

3.3. Performance metric evaluation in sales prediction

Evaluating the effectiveness of machine learning models in sales prediction is important to ensuring reliable and accurate forecasting. Various performance metrics are used in the literature to assess the predictive capabilities of different models, focusing primarily on error reduction, model interpretability, and adaptability to different datasets. The choice of evaluation metric depends on the nature of the prediction task, the dataset characteristics, and the business requirements of the study.

One of the most commonly used evaluation metrics is the Root Mean Squared Error (RMSE), which measures the average magnitude of prediction errors while penalizing larger deviations more than smaller ones. RMSE is widely adopted in sales forecasting studies across different sectors, including retail, e-commerce, and grocery sales, as it provides a clear understanding of the model's overall accuracy (Calitz et al., 2022; Chalapathy et al., 2022; Gupta et al., 2023; Kunal et al., 2024). For example, studies comparing regression-based models for store sales forecasting reported RMSE values for Random Forest and XGBoost, with XGBoost consistently achieving lower RMSE scores, indicating higher accuracy (Akanksha et al., 2022; Naik et al., 2022). Another frequently used metric is the Mean Absolute Error (MAE), which calculates the average absolute differences between predicted and actual sales values. Unlike RMSE, MAE treats all errors equally without squaring them, making it more interpretable in some business contexts (Behera & Nain, 2020; Li et al., 2021; Ratre & Jayaraj, 2023). Studies on Black Friday sales prediction have employed MAE alongside RMSE to compare the performance of models like Decision Tree Regression and Linear Regression, with results showing that ensemble methods like Random Forest and XGBoost tend to achieve lower MAE values (Huo, 2021; Rathore et al., 2024).

The Mean Squared Error (MSE) also used to evaluate model accuracy. Since MSE squares the errors, it emphasizes larger deviations more than RMSE does. Some studies prefer RMSE over MSE because RMSE maintains the same unit as the original sales data, making it easier to interpret (Behera & Nain, 2019; Kaushal et al., 2023; Nagaraj et al., 2023). For assessing the proportion of variance explained by the model, researchers use the R-squared (R^2) score. This metric indicates how well the independent variables explain the variance in the dependent variable (sales). A higher R^2 value suggests that the model captures more of the variation in sales trends. Studies in e-commerce and retail sales forecasting have found that XGBoost and ensemble models tend to achieve higher R^2 scores than traditional regression models, reinforcing their effectiveness in handling complex relationships between features (Liao et al., 2020; Nagaraj et al., 2023; Patil et al., 2023).

In time-series sales prediction studies, additional performance metrics such as the Root Mean Squared Scaled Error (RMSSE) and the F1 Score have been used. RMSSE is particularly relevant for sales forecasting applications that involve hierarchical time-series data, where the accuracy of forecasts at different levels (such as store-wise and category-wise sales) needs to be measured (Kamble et al., 2023; Loukili et al., 2024; Warpe et al., 2024). Meanwhile, the F1 Score, a metric commonly used in classification tasks, has been applied in ensemble-based models that classify sales trends or predict product demand categories (Pai et al., 2024; Singh et al., 2020; Sreemathy & Prasath, 2022).

Beyond traditional accuracy metrics, some studies emphasize the need to evaluate models based on computational efficiency and adaptability. The training time of models is often considered when comparing machine learning techniques, particularly for large datasets. Research on ensemble learning methods in retail sales forecasting has shown that while XGBoost and Random Forest achieve high accuracy, they require longer training times compared to simpler models like Linear Regression (Konda & Singh, 2023; Krishna et al., 2024; Li et al., 2023). Additionally, studies highlight that models should be tested for generalizability across different sales periods to ensure robustness in real-world applications (Chalapathy et al., 2022; Rekha et al., 2024; Shi & Zhang, 2023).

3.4. Sectors that faced challenges related to sales forcasting

Sales forecasting is critical across various sectors, as accurate predictions allow businesses to optimize inventory management, resource allocation, and strategic planning. However, different industries face distinct challenges due to the complexity of their sales patterns, external influencing factors, and data availability. Based on the analyzed literature, the retail industry, e-commerce, manufacturing, and the energy sector are among the most studied industries in sales forecasting, each encountering unique difficulties.

3.4.1. Retail industry

The retail industry is one of the most heavily researched sectors in sales forecasting due to its dependence on demand fluctuations, seasonality, and consumer behavior. Retailers struggle with demand volatility, which is influenced by promotions, holidays, economic conditions, and sudden shifts in consumer preferences (Bogaert et al., 2021; Manna et al., 2023; Spuritha et al., 2021). For example, Black Friday sales forecasting has been extensively studied, with research highlighting the difficulty of predicting consumer purchasing behavior due to massive price reductions and impulse buying (Rathore et al., 2024; Shymanskyi & Liaskovets, 2023; Singh et al., 2020). Furthermore, inventory management remains a major concern in retail, as inaccurate sales forecasts can lead to stockouts or overstocking, both of which negatively impact profitability (Behera & Nain, 2020; Cheriyan et al., 2018). There is another occurrence where businesses must forecast demand across online and offline stores. The rise of e-commerce has made it difficult for traditional retailers to predict demand accurately, as consumer purchasing behavior varies significantly between physical and digital platforms (Behera & Nain, 2019; Kamble et al., 2023). Some studies have proposed integrating customer segmentation techniques such as K-Means Clustering to enhance sales forecasting by identifying behavioral patterns among different customer groups (Calitz et al., 2022; Ratre & Jayaraj, 2023).

3.4.2. E-commerce sector

E-commerce platforms face challenges similar to those in retail but with additional complexities due to the dynamic nature of online sales, customer retention strategies, and personalization (Chalapathy et al., 2022; Huo, 2021; Shi & Zhang, 2023). Unlike traditional retail, e-commerce businesses must account for external factors such as website traffic, advertising effectiveness, and competitor pricing strategies, which significantly impact sales volume (Krishna et al., 2024; Latha et al., 2023; Sharma et al., 2019). Several studies have examined how machine learning models like XGBoost, Random Forest, and Neural Networks can handle large-scale transactional data to improve demand forecasting in online retail (Akanksha et al., 2022; Gunjal et al., 2022; Xavier et al., 2023). Major challenges in e-commerce are customer churn and retention, where forecasting sales is not just about predicting demand but also identifying which customers are likely to return for future purchases (Naik et al., 2022; Ramachandra et al., 2021) Studies have used clustering methods combined with regression models to predict better repeat purchases and long-term sales trends (Gupta et al., 2023; Pai et al., 2024).

3.4.3. Manufacturing and supply chain management

Manufacturing companies rely on sales forecasting to manage production planning, raw material procurement, and logistics. However, this sector faces challenges such as long supply chain lead times, fluctuating raw material costs, and demand uncertainty (Dairu & Shilong, 2021; Konda & Singh, 2023; Upadhyay et al., 2023). Inaccurate forecasts can lead to inefficiencies, including excess production or shortages that disrupt supply chains (Dutta et al., 2022; Tiganoaia & Anghel, 2024; Warpe et al., 2024). One of the biggest forecasting challenges in manufacturing is handling demand fluctuations caused by external shocks, such as geopolitical events, economic downturns, and pandemics. Research has explored how time-series forecasting models and ensemble learning techniques can improve prediction accuracy in volatile markets (Han & Li, 2024; Sadasivam et al., 2024; Samara & Stanich, 2023). Additionally, multivariate forecasting models that incorporate economic indicators, market trends, and production constraints have been proposed to capture better the complex factors affecting sales in manufacturing (Nagaraj et al., 2023; Nebri et al., 2024).

3.4.4. Energy and utility sector

Sales forecasting in the energy sector is crucial for balancing supply and demand, particularly in industries such as electricity and gas distribution. Unlike retail and manufacturing, where sales data is primarily driven by customer demand and pricing strategies, energy sales forecasting must account for weather conditions, consumption patterns, and government regulations (Ashraf et al., 2024; Harshini et al., 2024; Saito et al., 2021). For example, studies on electricity sales prediction highlight the difficulty of accurately modeling consumption due to fluctuations in temperature, renewable energy integration, and unpredictable demand spikes (Merdas & Mousa, 2023; Rekha et al., 2024; Wang, 2020). Energy and utility sales also struggle with forecasting data granularity and availability. Some studies emphasize the need for high-frequency data collection and real-time forecasting methods to improve predictive accuracy (Bohanec et al., 2017; Li et al., 2021; Mondal et al., 2024). Machine learning techniques such as LSTM networks and hybrid models combining statistical and deep learning approaches have been explored to address these challenges (Ajaykrishna et al., 2023; Liang et al., 2024; Wisesa et al., 2020a).

3.5. How to distinguish real problem-based research

In sales forecasting, there is a clear distinction between research that is problem-driven and research that primarily focuses on improving predictive models. Problem-driven research aims to solve industry-specific challenges, considering external factors such as economic conditions, consumer behavior, and supply chain constraints, whereas model-focused research primarily optimizes algorithms for performance metrics without necessarily addressing real-world complexities (Behera & Nain, 2019; Fourkiotis & Tsadiras, 2024; Latha et al., 2023; Wisesa et al., 2020a).

One key differentiator is the inclusion of external variables beyond historical sales data. Studies that address real-world problems often incorporate economic indicators, customer demographics, marketing strategies, or supply chain logistics. For example, research on sales prediction in the retail sector has explored the impact of seasonal trends, promotional campaigns, and consumer sentiment analysis to enhance forecast accuracy (Behera & Nain, 2020; Merdas & Mousa, 2023; Mohamed-Amine et al., 2024; Spuritha et al., 2021). Similarly, studies in manufacturing and supply chain management integrate supplier reliability and production constraints to ensure inventory optimization aligns with demand forecasting (Dairu & Shilong, 2021; Han & Li, 2024; Upadhyay et al., 2023). In contrast, studies that focus solely on refining models, such as optimizing hyperparameters of machine learning algorithms, may lack practical implementation details (Ashraf et al., 2024; Gunjal et al., 2022).

Another major distinction is the extent of industry collaboration and real-world validation. Studies that work closely with businesses, utilizing proprietary datasets and integrating domain expertise, tend to provide more actionable insights. For instance, e-commerce sales prediction research often collaborates with online platforms to incorporate real-time pricing and customer behavior data, leading to solutions that can be directly applied in business strategies (Branda et al., 2020; Gupta et al., 2023; Latha et al., 2023). On the other hand, many benchmark-driven studies rely solely on open datasets like Kaggle, which, while useful for model development, may not capture real-world complexities such as data drift and market disruptions (Pai et al., 2024; Sharma et al., 2019).

Problem-solving research often evaluates model performance beyond standard accuracy metrics by considering business impact. While RMSE and R-squared are widely used, studies addressing real problems may also assess cost savings, revenue growth, inventory efficiency, or customer satisfaction improvements. For example, research in the energy sector evaluates forecasting models not just based on numerical accuracy but also on their ability to prevent resource shortages and optimize distribution (Bohanec et al., 2017; Kunal et al., 2024; Wisesa et al., 2020b). The adaptability of research findings to practical applications is a key criterion. Studies that propose actionable recommendations for decision-makers, such as pricing strategies, marketing optimizations, or inventory adjustments, demonstrate higher real-world relevance. In contrast, research that strictly benchmarks models without discussing implementation feasibility often lacks practical utility (Bogaert et al., 2021; Saito et al., 2021; Wang, 2020).

4. Final Considerations

This systematic review examined the diverse range of machine learning techniques applied to sales forecasting and the various performance metrics used to evaluate these models. The analysis revealed that traditional regression models, such as linear and decision tree regression, continue to serve as important baselines; however, ensemble methods like random forest and XGBoost have emerged as preferred approaches due to their ability to capture complex, nonlinear relationships and deliver higher accuracy. Deep learning models, including LSTM networks and artificial neural networks, have also been explored, particularly for time-series forecasting, though their practical application may be limited by computational demands and data requirements.

The evaluation of these models predominantly relies on error-based metrics such as RMSE, MAE, MSE, and R-squared, with additional measures like MAPE and RMSSE used in specific contexts. These metrics provide critical insights into the predictive performance and variance explanation capabilities of the models. Moreover, the review highlighted that different sectors face unique challenges in sales forecasting. For instance, physical retail and fashion industries grapple with demand volatility and seasonal fluctuations, e-commerce platforms must contend with rapidly shifting consumer behavior and online market dynamics, while manufacturing and supply chain management deal with long lead times and external economic shocks. The energy and utility sectors, on the other hand, require models that can accommodate weather conditions, regulatory changes, and high-frequency data.

A further distinction emerged between research that addresses real-world problems and studies focused solely on model optimization. Research with a problem-driven approach typically integrates external variables such as economic indicators, customer demographics, and operational constraints, and often involves close collaboration with industry partners to ensure practical applicability. In contrast, model-focused research tends to rely on benchmark datasets and standard performance metrics without fully considering the complex, dynamic nature of real-world sales environments.

Overall, the literature marks the importance of selecting machine learning models that align with the specific data characteristics and business contexts of each industry. Future research should focus on developing hybrid and adaptive models that not only enhance predictive accuracy but also provide actionable insights for real-world decision-making in sales forecasting.

Acknowledgment

The authors would like to thank the Faculty of Engineering, Diponegoro University, Indonesia, which provided financial support for this research. The authors also appreciate the help and cooperation of the participants who participated in this study.

Ethical Considerations

Not applicable.

Conflict of Interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Achumie, G. O., Oyegbade, I. K., Igwe, A. N., Ofodile, O. C., & Azubuike, C. (2022). AI-driven predictive analytics model for strategic business development and market growth in competitive industries. *International Journal of Social Science Exceptional Research, 1*(1), 13–25. https://doi.org/10.54660/IJSSER.2022.1.1.13-25

Ajaykrishna, S., Suganya, T. S., Rao, B., & Pughazendi, N. (2023). Online sales prediction in e-commerce market using machine learning. In *Proceedings of the 4th International Conference on Signal Processing and Communication* (pp. 47–51). https://doi.org/10.1109/ICSPC57692.2023.10125855

Akanksha, A., Yadav, D., Jaiswal, D., Ashwani, A., & Mishra, A. (2022). Store-sales forecasting model to determine inventory stock levels using machine learning. In *Proceedings of the 5th International Conference on Inventive Computation Technologies* (pp. 339–344). https://doi.org/10.1109/ICICT54344.2022.9850468

Akusta, A. (2024). Comparative evaluation of statistical models and machine learning approaches in modelling the energy dependency of the BIST industrial index: Balancing predictive performance and interpretability. *Journal of Corporate Governance, Insurance, and Risk Management, 11*(3), 160–167. https://doi.org/10.56578/jcgirm110302

Ali, R. F., Muneer, A., Almaghthawi, A., Alghamdi, A., Fati, S. M., & Ghaleb, E. A. A. (2023). BMSP-ML: Big mart sales prediction using different machine learning techniques. *International Journal of Artificial Intelligence, 12*(2), 874–883. https://doi.org/10.11591/ijai.v12.i2.pp874-883

Ashraf, K., Lee, K., Kim, G., & Kang, J. Y. (2024). Sales in commercial alleys and their association with air pollution: Case study in South Korea. *Sustainability (Switzerland), 16*(2), 1–19. https://doi.org/10.3390/su16020530

Behera, G., & Nain, N. (2019). Grid search optimization (GSO) based future sales prediction for big mart. In *Proceedings of the 15th International Conference on Signal Image Technology and Internet Based Systems* (pp. 172–178). https://doi.org/10.1109/SITIS.2019.00038

Behera, G., & Nain, N. (2020). A comparative study of big mart sales prediction. In *Communications in Computer and Information Science* (Vol. 1147, pp. 421–432). https://doi.org/10.1007/978-981-15-4015-8_37

Bogaert, M., Ballings, M., Van den Poel, D., & Oztekin, A. (2021). Box office sales and social media: A cross-platform comparison of predictive ability and mechanisms. *Decision Support Systems, 147*, 113517. https://doi.org/10.1016/j.dss.2021.113517

Bohanec, M., Robnik-Šikonja, M., & Kljajić Borštnar, M. (2017). Organizational learning supported by machine learning models coupled with general explanation methods: A case of B2B sales forecasting. *Organizacija, 50*(3), 217–233. https://doi.org/10.1515/orga-2017-0020

Branda, F., Marozzo, F., & Talia, D. (2020). Ticket sales prediction and dynamic pricing strategies in public transport. *Big Data and Cognitive Computing, 4*(4), 36. https://doi.org/10.3390/bdcc4040036

Calitz, A. P., Cullen, M., & Mamombe, S. (2022). Predictive analysis of fuel prices using machine learning. In *Proceedings of the 3rd International Conference on Next Generation Computing Applications (NextComp 2022)* (pp. 1–6). https://doi.org/10.1109/NextComp55567.2022.9932204

Chalapathy, N., Helen Josephine, V. L., & Rajalakshmi, K. (2022). Sales prediction scheme using RFM-based clustering and regressor model for e-commerce company. In *Proceedings of the ACM International Conference Proceeding Series* (pp. 1–6). https://doi.org/10.1145/3590837.3590937

Cheriyan, S., Ibrahim, S., Mohanan, S., & Treesa, S. (2018). Intelligent sales prediction using machine learning techniques. In *Proceedings of the International Conference on Computing, Electronics and Communications Engineering* (pp. 53–58). https://doi.org/10.1109/iCCECOME.2018.8659115

Ching, T., Himmelstein, D. S., Beaulieu-Jones, B. K., Kalinin, A. A., Do, B. T., Way, G. P., Ferrero, E., Agapow, P. M., Zietz, M., Hoffman, M. M., Xie, W., Rosen, G. L., Lengerich, B. J., Israeli, J., Lanchantin, J., Woloszynek, S., Carpenter, A. E., Shrikumar, A., Xu, J., … Greene, C. S. (2018). Opportunities and obstacles for deep learning in biology and medicine. *Journal of the Royal Society Interface, 15*(141), 20170387. https://doi.org/10.1098/rsif.2017.0387

Dairu, X., & Shilong, Z. (2021). Machine learning model for sales forecasting by using XGBoost. In *Proceedings of the IEEE International Conference on Consumer Electronics and Computer Engineering* (pp. 480–483). https://doi.org/10.1109/ICCECE51280.2021.9342304

Dritsas, E., & Trigka, M. (2025). Exploring the intersection of machine learning and big data: A survey. *Machine Learning and Knowledge Extraction, 7*(1), 13. https://doi.org/10.3390/make7010013

Dutta, S. R., Das, S., & Chatterjee, P. (2022). Smart sales prediction of pharmaceutical products. In *Proceedings of the 8th International Conference on Smart Structures and Systems* (pp. 1–6). https://doi.org/10.1109/ICSSS54381.2022.9782271

Elhanashi, A., Dini, P., Saponara, S., & Zheng, Q. (2023). Integration of deep learning into the IoT: A survey of techniques and challenges for real-world applications. *Electronics (Switzerland), 12*(24), 4935. https://doi.org/10.3390/electronics12244925

Fotache, M., Cojocariu, I. C., & Bertea, A. (2021). High-level machine learning framework for sports events ticket sales prediction. In *Proceedings of the ACM International Conference Proceeding Series* (pp. 55–60). https://doi.org/10.1145/3472410.3472426

Fourkiotis, K. P., & Tsadiras, A. (2024). Applying machine learning and statistical forecasting methods for enhancing pharmaceutical sales predictions. *Forecasting, 6*(1), 170–186. https://doi.org/10.3390/forecast6010010

Gunjal, S. N., Kshirsagar, D. B., Dange, B. J., & Khodke, H. E. (2022). Fusing clustering and machine learning techniques for Big-Mart sales predication. In *Proceedings of the IEEE International Conference on Blockchain and Distributed Systems Security* (pp. 1–6). https://doi.org/10.1109/ICBDS53701.2022.9935906

Gupta, G., Gupta, K. L., & Kansal, G. (2023). MegaMart sales prediction using machine learning techniques. In *Proceedings of Third International Conference on Computing, Communications* (pp. 437–446). https://doi.org/10.1007/978-981-19-1142-2_35

Han, B., & Li, P. (2024). Apparel livestreaming sales forecasting models based on machine learning algorithms. *Journal of Silk, 61*(7), 109–117. https://doi.org/10.3969/j.issn.1001-7003.2024.07.012

Harshini, G., Reddy, T. T., Prathima, T., & Sirisha, A. (2024). Automated product categorization and sales forecasting: A machine learning approach for retail analytics. *Grenze International Journal of Engineering & Technology (GIJET), 10*(2), 924–930.

Huo, Z. (2021). Sales prediction based on machine learning. In *Proceedings of the 2nd International Conference on E-Commerce and Internet Technology* (pp. 410–415). https://doi.org/10.1109/ECIT52743.2021.00093

Ibrahim, N. F., & Wang, X. (2019). Decoding the sentiment dynamics of online retailing customers: Time series analysis of social media. *Computers in Human Behavior, 96,* 32–45. https://doi.org/10.1016/j.chb.2019.02.004

Jun, S. P., Park, D. H., & Yeom, J. (2014). The possibility of using search traffic information to explore consumer product attitudes and forecast consumer preference. *Technological Forecasting and Social Change, 86,* 237–253. https://doi.org/10.1016/j.techfore.2013.10.021

Kamble, T., Ghuge, M., Rana, R., Vardhan, H., Shelar, Y., & Machale, T. (2023). Ensemble machine learning models to forecast sales. In *Proceedings of the 3rd International Conference on Innovative Mechanisms for Industry Applications* (pp. 1056–1061). https://doi.org/10.1109/ICIMIA60377.2023.10426565

Kaushal, R. K., Giri, K. K. B., Kumar, H. S., Raina, K., Choudhury, D., & Naikade, K. (2023). AI-based approach for retail sale forecasting. In *Proceedings of the International Conference on Self Sustainable Artificial Intelligence Systems* (pp. 111–116). https://doi.org/10.1109/ICSSAS57918.2023.10331651

Konda, S., & Singh, S. P. (2023). An in-depth evaluation of machine learning techniques for anticipating effective human health outcomes. In *Proceedings of the International Conference on Artificial Intelligence for Innovations in Healthcare Industries* (pp. 1–6). https://doi.org/10.1109/ICAIIHI57871.2023.10489226

Krishna, R. S., Deepa, K., & Sangeetha, S. V. T. (2024). Performance analysis of ML techniques in EV sales – A study. In *Proceedings of the 10th International Conference on Advanced Computing and Communication Systems* (pp. 1045–1049). https://doi.org/10.1109/ICACCS60874.2024.10717024

Kunal, Singh, M., Sharma, H., Roy, S., & Agarwal, K. (2024). Big Mart sales analysis. In *Proceedings of the 6th International Conference on Computational Intelligence and Communication Technologies* (pp. 554–561). https://doi.org/10.1109/CCICT62777.2024.00092

Latha, S. B., Dastagiraiah, C., Kiran, A., Asif, S., Elangovan, D., & Reddy, P. C. S. (2023). An adaptive machine learning model for Walmart sales prediction. In *Proceedings of the International Conference on Circuit Power and Computing Technologies* (pp. 988–992). https://doi.org/10.1109/ICCPCT58313.2023.10245029

Li, J., Zhang, Z., Kang, S., Lu, Y., Gan, L., & Xu, Y. (2023). Research on future sales prediction based on feature engineering. In *Proceedings of the ACM International Conference Proceeding Series* (pp. 223–228). https://doi.org/10.1145/3652628.3652666

Li, J., Zheng, Y., Hu, H., Lu, J., & Zhan, C. (2021). Predicting video game sales based on machine learning and hybrid feature selection method. In *Proceedings of the IEEE International Conference on Intelligent Systems and Knowledge Engineering* (pp. 497–502). https://doi.org/10.1109/ISKE54062.2021.9755343

Liang, X.-X., Lin, Y., Deng, C., Mo, Y., Lu, B., Yang, J., Xu, L., Li, P., Liu, X., Chen, K., & Liu, M. (2024). Machine learning-based sales prediction using Bayesian optimized XGBoost algorithms. In *Proceedings of the Modern Management Based on Big Data V* (pp. 248–268). https://doi.org/10.3233/faia240263

Liao, W., Ye, G., Yin, Y., Yan, W., Ma, Y., & Zuo, D. (2020). Auto parts sales prediction based on machine learning for small data and a long replacement cycle. In *Proceedings of the IEEE International Conference on Computer Systems and Applications* (pp. 1–5). https://doi.org/10.1109/AICCSA50499.2020.9316540

Lin, Y. C., Sung, B., & Park, S. Do. (2024). Integrated systematic framework for forecasting China's consumer confidence: A machine learning approach. *Systems, 12*(11), 1–33. https://doi.org/10.3390/systems12110445

Loukili, M., Messaoudi, F., Ghazi, M. El, & Azirar, H. (2024). Predicting future sales: A machine learning algorithm showdown. In *Lecture Notes in Networks and Systems* (Vol. 837, pp. 26–31). https://doi.org/10.1007/978-3-031-48465-0_4

Madhuvanthi, K., Kailasanathan, N., C, S. N., & Somayaji, S. (2019). Car sales prediction using machine learning algorithms. *International Journal of Innovative Technology and Exploring Engineering, 8*(5), 2278–3075. https://www.researchgate.net/publication/332072545

Mallik, R. S., Abhiram, R., Reddy, S. R., & Jagadish, R. M. (2022). A comprehensive survey on sales forecasting models using machine learning algorithms. In *Proceedings of the 4th International Conference on Emerging Research in Electronics, Computer Science and Technology* (pp. 1–6). https://doi.org/10.1109/ICERECT56837.2022.10060168

Manna, A., Kolpe, K., Mhalungekar, A., Pattewar, S., Kaloge, P., & Patil, R. (2023). Comparative study of various machine learning algorithms for pharmaceutical drug sales prediction. In *Proceedings of 2023 International Conference on Signal Processing, Computation, Electronics, Power and Telecommunication (IConSCEPT 2023)* (pp. 1–6). https://doi.org/10.1109/IConSCEPT57958.2023.10169964

Merdas, H. M., & Mousa, A. H. (2023). Food sales prediction model using machine learning techniques. *International Journal of Electrical and Computer Engineering, 13*(6), 6578–6585. https://doi.org/10.11591/ijece.v13i6.pp6578-6585

Mohamed-Amine, N., Abdellatif, M., & Belaid, B. (2024). Artificial intelligence for forecasting sales of agricultural products: A case study of a Moroccan agricultural company. *Journal of Open Innovation: Technology, Market, and Complexity, 10*(1), 1–10. https://doi.org/10.1016/j.joitmc.2023.100189

Mondal, S., Debbarma, A., & Prakash, B. (2024). Big Mart sales prediction using machine learning. In *Proceedings of the 2024 10th International Conference on Communication and Signal Processing (ICCSP 2024)* (pp. 742–747). https://doi.org/10.1109/ICCSP60870.2024.10544274

Nagaraj, P., Nani, K., Krishna, E. T., Reddy, K. A. K., Sekar, R. R., & Rajkumar, T. D. (2023). Customer sale analysis and classification using machine learning algorithm. In *Proceedings of the International Conference on Data Science, Agents and Artificial Intelligence* (pp. 1–5). https://doi.org/10.1109/ICDSAAI59313.2023.10452665

Naik, H., Yashwanth, K., Suraj, P., & Jayapandian, N. (2022). Machine learning-based food sales prediction using random forest regression. In *Proceedings of the 6th International Conference on Electronics, Communication and Aerospace Technology* (pp. 998–1004). https://doi.org/10.1109/ICECA55336.2022.10009277

Nebri, M. A., Moussaid, A., & Bouikhalene, B. (2024). Forecasting livestock feed sales using machine learning techniques: An analysis of the Moroccan market. *Indonesian Journal of Electrical Engineering and Computer Science, 35*(2), 1139–1150. https://doi.org/10.11591/ijeecs.v35.i2.pp1139-1150

Niu, Y. (2020). Walmart sales forecasting using XGBoost algorithm and feature engineering. In *Proceedings of the International Conference on Big Data and Artificial Intelligence and Software Engineering* (pp. 458–461). https://doi.org/10.1109/ICBASE51474.2020.00103

Pai, K., Batool, M., Ravi, N., Surendra, R., Shree, A. N. R., & Meenakshi, S. J. (2024). Enhancing sales forecasting and prediction with cutting-edge machine learning methods. In *Proceedings of the 2nd IEEE International Conference on Data Science and Information System* (p. 6). https://doi.org/10.1109/ICDSIS61070.2024.10594627

Patil, S., Nankar, O., Agrawal, R., Sharma, K., Awasthi, S., & Jha, N. (2023). Black Friday sales prediction using supervised machine learning. In *Proceedings of the International Conference on Artificial Intelligence and Smart Communication* (pp. 1006–1012). https://doi.org/10.1109/AISC56616.2023.10084959

Ramachandra, H. V., Balaraju, G., Rajashekar, A., & Patil, H. (2021). Machine learning application for Black Friday sales prediction framework. In *Proceedings of the International Conference on Emerging Smart Computing and Informatics* (pp. 57–61). https://doi.org/10.1109/ESCI50559.2021.9396994

Rane, N., Choudhary, S. P., & Rane, J. (2024). Ensemble deep learning and machine learning: Applications, opportunities, challenges, and future directions. *Studies in Medical and Health Sciences, 1*(2), 18–41. https://doi.org/10.48185/smhs.v1i2.1225

Rathore, S. P. S., Bihade, V. M., Sane, S., Limbore, N. V., Lenka, R. M., & Priyanka. (2024). Smart model for Black Friday sales prediction. In *Proceedings of the International Conference on Smart Devices* (pp. 1–5). https://doi.org/10.1109/ICSD60021.2024.10751174

Ratre, S., & Jayaraj, J. (2023). Sales prediction using ARIMA, Facebook's Prophet and XGBoost model of machine learning. In *Proceedings of the Machine Learning, Image Processing, Network Security and Data Sciences* (pp. 101–111). https://doi.org/10.1007/978-981-19-5868-7_9

Rekha, C. S., Sathwika, K. B., Nandini, G. S., Sumanth, R., & Shareefunnisa, S. (2024). Ensemble strategies for predicting Walmart sales using regression models. In *Proceedings of the 8th International Conference on Inventive Systems and Control* (pp. 693–699). https://doi.org/10.1109/ICISC62624.2024.00123

Rozony, F. Z., Aktar, M. N. A., Ashrafuzzaman, M., & Islam, A. (2024). A systematic review of big data integration challenges and solutions for heterogeneous data sources. *Academic Journal on Business Administration, Innovation & Sustainability, 4*(4), 1–18. https://doi.org/10.69593/ajbais.v4i04.111

Sadasivam, V. R., Narendharakumar, G., Ragavan, M., & Prabhuram, A. (2024). Food sales analysis and prediction using machine learning. In *Proceedings of the 2nd International Conference on Artificial Intelligence and Machine Learning Applications: Healthcare and Internet of Things* (pp. 1–6). https://doi.org/10.1109/AIMLA59606.2024.10531342

Saito, F., Yamashita, H., Sasaki, H., & Goto, M. (2021). Machine learning for demand prediction of seasonal second-hand fashion items based on prior and fine-tuning prediction models. In *Proceedings of the IEEE 12th International Workshop on Computational Intelligence and Applications* (pp. 1–6). https://doi.org/10.1109/IWCIA52852.2021.9626037

Samara, K., & Stanich, M. (2023). Using machine learning to predict grocery sales. In *Proceedings of Intelligent Computing* (pp. 936–943). https://doi.org/10.1007/978-3-031-37963-5_65

Sharma, S. K., Chakraborti, S., & Jha, T. (2019). Analysis of book sales prediction at Amazon marketplace in India: A machine learning approach. *Information Systems and E-Business Management, 17*(2–4), 261–284. https://doi.org/10.1007/s10257-019-00438-3

Shi, R., & Zhang, C. (2023). A study of sales forecasting in multinational retail companies: A feature extraction–machine learning–classification based forecasting framework. In *Proceedings of the IEEE International Conference on Sensors, Electronics and Computer Engineering* (pp. 401–405). https://doi.org/10.1109/ICSECE58870.2023.10263406

Shymanskyi, V., & Liaskovets, V. (2023). Cascade model for price and time of car sales prediction. In *CEUR Workshop Proceedings* (Vol. 3641, pp. 152–167). https://ceur-ws.org/Vol-3641/paper14.pdf

Singh, K., Booma, P. M., & Eaganathan, U. (2020). E-commerce system for sale prediction using machine learning technique. In *Journal of Physics: Conference Series, 1712*(1), 1–9. https://doi.org/10.1088/1742-6596/1712/1/012042

Spuritha, M., Kashyap, C. S., Nambiar, T. R., Kiran, D. R., Rao, N. S., & Reddy, G. P. (2021). Quotidian sales forecasting using machine learning. In *Proceedings of the IEEE International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems* (pp. 1–6). https://doi.org/10.1109/ICSES52305.2021.9633975

Sreemathy, J., & Prasath, N. (2022). Machine learning-based sales prediction and characterization using consumer perceptions. In *Proceedings of the 6th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)* (pp. 1072–1076). https://doi.org/10.1109/I-SMAC55078.2022.9987359

Tiganoaia, B., & Anghel, I.-P. (2024). Machine learning algorithms for sales prediction in stores – An applied research. *U.P.B. Scientific Bulletin, Series C, 86*(2). https://nrf.com/media-center/press-releases/thanksgiving-draws-nearly-190-million-shoppers

Upadhyay, H., Shekhar, S., Vidyarthi, A., Prakash, R., & Gowri, R. (2023). Sales prediction in the retail industry using machine learning: A case study of BigMart. In *Proceedings of the IEEE International Conference on Electrical, Electronics, Communication and Computers* (pp. 1–6). https://doi.org/10.1109/ELEXCOM58812.2023.10370313

Wang, J. (2020). A hybrid machine learning model for sales prediction. In *Proceedings of the International Conference on Intelligent Computing and Human-Computer Interaction* (pp. 363–366). https://doi.org/10.1109/ICHCI51889.2020.00083

Warpe, V. S., Buchkul, S. D., Chobe, P., & Pardeshi, D. B. (2024). EV sales price forecasting using machine learning. In *Proceedings of the 2nd International Conference on Sustainable Computing and Smart Systems* (pp. 818–823). https://doi.org/10.1109/ICSCSS60660.2024.10625450

Wen, K. Y., Joseph, M. H., & Sivakumar, V. (2024). Big Mart sales prediction using machine learning. *EAI Endorsed Transactions on Internet of Things, 10,* 1–6. https://doi.org/10.4108/eetiot.6453

Wisesa, O., Adriansyah, A., & Khalaf, O. I. (2020a). Prediction analysis for business-to-business (B2B) sales of telecommunication services using machine learning techniques. *Majlesi Journal of Electrical Engineering, 14*(4), 145–153. https://doi.org/10.29252/mjee.14.4.145

Wisesa, O., Adriansyah, A., & Khalaf, O. I. (2020b). Prediction analysis sales for corporate services telecommunications company using gradient boost algorithm. In *Proceedings of the 2nd International Conference on Broadband Communications, Wireless Sensors and Powering* (pp. 101–106). https://doi.org/10.1109/BCWSP50066.2020.9249397

Xavier, A. K., Manjunath, C., Manohar, M., Gurudas, V. R., Jayapandian, N., & Balamurugan, M. (2023). Analytical methods of machine learning model for e-commerce sales analysis and prediction. In *Proceedings of IEEE InC4 2023 – 2023 IEEE International Conference on Contemporary Computing and Communications.* https://doi.org/10.1109/InC457730.2023.10263003