Journal of Computer Science

Research Article

Predictive Mathematical Modeling and Classification of Retail Sales Orders Using AI Machine Learning Techniques

Mohammad Subhi Al-Batah[1], Mowafaq Salem Alzboon[1] and Hamzeh Zureigat[2]

[1]*Department of Computer Science, Faculty of Information Technology, Jadara University, Irbid, Jordan*
[2]*Department of Mathematics, Faculty of Science and Technology, Jadara University, Irbid 21110, Jordan*

Article history
Received: 22-07-2025
Revised: 03-12-2025
Accepted: 04-12-2025

Corresponding Author:
Mohammed Subhi Al-batah
Department of Computer Science, Jadara University, Irbid, Jordan
Email: albatah@jadara.edu.jo

Abstract: This study presents a systematic mathematical model known as a low-code approach to classifying retail sales orders by size using AI machine learning techniques within the Orange Data Mining platform. Leveraging a real-world sales dataset sourced from Kaggle, we implemented and evaluated ten classification models, including ensemble learners (AdaBoost, Gradient Boosting), probabilistic classifiers (Naïve Bayes), distance-based methods (kNN), and interpretable algorithms (CN2 Rule Induction). Each model was assessed through 10-fold cross-validation using performance metrics such as accuracy, F1-score, precision, recall, AUC, and LogLoss. The experimental workflow integrated visual preprocessing, model training, and comparative evaluation, enabling reproducibility without programming expertise. The results reveal that ensemble models, particularly AdaBoost, achieved perfect classification accuracy (100%) and AUC (1.000), while CN2 Rule Induction offered near-perfect accuracy (99.8%) alongside interpretable rule-based outputs. Traditional models like Logistic Regression and kNN also demonstrated strong performance but were outperformed by advanced ensembles. This research contributes a novel combination of high-performing and explainable models in a retail classification task using a low-code framework. The proposed approach provides practical guidance for retailers, analysts, and educators seeking accurate and accessible predictive tools for operational decision-making. Future directions include multi-class extension, imbalance handling, and deployment in real-time environments.

Keywords: Predictive Mathematical Modeling, Sales Classification, AI Machine Learning, Orange Data Mining, Retail Analytics

Introduction

In today's data-driven economy, retail enterprises generate vast amounts of transactional data that, when analyzed effectively, can offer valuable insights for strategic decision-making. One critical aspect of retail analytics is the classification of sales orders, which helps businesses optimize supply chain operations, forecast demand, and personalize customer engagement. Traditional rule-based systems often fall short when dealing with complex patterns and nonlinear interactions among features. This limitation has prompted a shift toward data-driven predictive modeling using Machine Learning (ML) techniques, which excel at uncovering hidden structures in large datasets.

Despite the proliferation of ML applications across diverse sectors, including healthcare, finance, and manufacturing, retail classification tasks remain underexplored, especially those that involve nuanced distinctions such as order size categorization (Al-Batah et al., 2018). Prior studies have predominantly focused on high-level metrics such as total sales or customer segmentation, with limited attention to order-level classification using end-to-end machine learning pipelines (Mistry and Savani, 2022).

The motivation behind this study stems from the practical need to automate and enhance the classification of sales order sizes (e.g., "Small" vs. "Medium") using transaction-level data. Accurately classifying such orders is vital for improving inventory management, minimizing delivery delays, and ensuring responsive customer service. This is particularly relevant for e-commerce businesses and omnichannel retailers, where rapid decision-making is essential.

The key objectives of this research are as follows:

(i) To evaluate and compare the effectiveness of diverse machine learning models including ensemble, probabilistic, and rule-based algorithms in classifying retail sales orders
(ii) To leverage Orange Data Mining as a low-code tool for enabling reproducible and accessible experimentation
(iii) To interpret model results through confusion matrices, performance metrics (e.g., AUC, F1 score, CA), and graphical analyses (ROC curves), offering practical guidance for model deployment in retail contexts

This paper contributes to the literature by providing a comprehensive empirical evaluation of classification models on a real-world retail dataset sourced from Kaggle, using Orange Data Mining for implementation. Unlike previous studies that focused solely on accuracy, this work investigates a wide range of evaluation metrics, including LogLoss and model interpretability. Furthermore, it highlights the performance and reliability of ensemble methods such as AdaBoost and Gradient Boosting, which despite their popularity have not been widely benchmarked in retail classification problems involving small-to-medium order differentiation (Liang et al., 2023; Prasetyo and Sari, 2020).

The novelty of this research lies in combining interpretability (via CN2 Rule Induction) and performance (via ensemble methods) within a unified low-code workflow. This makes the approach replicable for practitioners without deep programming expertise, thus bridging the gap between academic modeling and real-world business applications.

Literature Review

The use of Machine Learning (ML) in retail analytics has garnered significant attention in recent years, enabling businesses to gain deeper insights from transactional data. Several researchers have explored classification-based approaches for retail applications, yet few have focused on the specific task of order size classification using low-code tools like Orange. This section is structured into three key themes: Classification algorithms in retail datasets, ensemble learning techniques, and the emergence of low-code/no-code platforms for ML applications.

Classification Techniques in Retail Data

Traditional classifiers such as logistic regression, decision trees, and Support Vector Machines (SVM) have long been utilized in retail environments due to their simplicity and predictive power. Qureshi and Qureshi (2021) used decision tree models in classifying retail transaction records, achieving moderate performance in binary classification problems. Kumar et al. (2022) applied SVM to categorize products in supermarket data and noted high performance but limited interpretability in business settings. Similarly, Sharma and Thakur (2023) highlighted that probabilistic models like Naïve Bayes and instance-based learners like k-Nearest Neighbors (kNN) can provide effective classification under certain conditions, though they are sensitive to data distributions and scalability.

Ensemble Learning in Retail Classification

Recent studies have demonstrated that ensemble methods such as Random Forest, AdaBoost, and Gradient Boosting generally outperform individual classifiers in terms of robustness and accuracy. Tripathy et al. (2022) reported that Random Forest significantly enhanced the prediction of customer preferences in retail CRM datasets. Das and Sen (2023) evaluated AdaBoost models for customer purchase prediction in online retail and found the ensemble technique superior to traditional methods. Osei-Bryson and Liu (2021) also noted that ensemble models improve classification stability in churn prediction tasks. However, these models have seldom been applied to distinguish subtle class differences such as "Small" vs. "Medium" orders in a retail dataset.

Low-Code Tools in Applied Machine Learning

The democratization of machine learning has led to the emergence of low-code environments such as Orange, KNIME, and RapidMiner (Al-Batah, 2025). These platforms allow non-programmers to perform complex data analysis using graphical interfaces. Hassan et al. (2023) explored the use of Orange for credit risk classification in banking and confirmed the utility of visual workflows for rapid prototyping. Nonetheless, limited work exists on applying Orange to retail datasets, particularly for multi-model comparison and benchmarking. Moreover, rule-based classifiers like CN2 Rule Induction known for their interpretability have not been widely benchmarked against ensemble models in retail applications using low-code tools.

Research Gap and Contribution

From the reviewed literature, it is evident that previous studies typically focus on either single-model evaluation or ensemble comparisons in narrow domains like churn or credit risk. There is a lack of studies that:

(i) Apply a comprehensive range of classifiers including ensemble, probabilistic, and rule-based models on retail order size classification tasks
(ii) Use low-code tools to facilitate deployment and experimentation
(iii) Emphasize diverse evaluation metrics beyond accuracy, such as F1 score, LogLoss, and model interpretability

This study addresses these gaps by implementing a low-code comparative analysis on retail sales data using Orange Data Mining, evaluating ten classification algorithms, and assessing both performance and interpretability aspects.

Materials

- Computing Environment: Hardware - (e.g., personal laptop, specifications)
- Software Versions: Orange Data Mining v3.29, Python 3.8, etc
- Parameter Settings: All models were set to default parameters unless otherwise specified

Methods

This section outlines the dataset used in the study, the machine learning models employed, the experimental workflow, and the evaluation strategy. The entire process was implemented using the Orange Data Mining platform to ensure reproducibility, clarity, and accessibility.

Dataset Description

The dataset employed in this research is titled "Sample Sales Data", publicly available via Kaggle (Kyanyoga, 2022).

This dataset simulates a retail environment, consisting of 2,823 sales transaction records. Each record captures various attributes related to sales, such as order ID, customer name, product type, shipping method, order priority, unit price, and the target variable: Order Size, which is labeled either *Small* or *Medium*. The data structure is summarized in Table 1.

The dataset was small (2,823 records) and simulated, which limits real-world generalizability. This should be considered when interpreting the results. Statistical validation, such as t-tests or ANOVA, was conducted to assess the significance of model performance differences.

Additional fields include unit price, shipping cost, order date, and delivery mode.

Experimental Workflow

The predictive modeling process was executed in five primary phases:

1. Data Acquisition
2. The dataset is imported into Orange from CSV format.
3. Data Preprocessing

- o Missing values are identified and imputed.
- o Categorical variables are transformed into numerical representations using one-hot encoding.
- o Feature selection is performed to retain the most informative predictors.

4. Model Training
5. Ten machine learning classifiers were trained

- o Logistic Regression
- o k-Nearest Neighbors (kNN)
- o Naïve Bayes
- o Random Forest
- o Support Vector Machine (SVM)
- o Neural Network
- o AdaBoost
- o Gradient Boosting
- o CN2 Rule Induction
- o Constant (Baseline)

6. Model Evaluation:
7. Each model was evaluated using 10-fold cross-validation to ensure statistical robustness. The metrics calculated include

- o Area Under Curve (AUC)
- o Classification Accuracy (CA)
- o F1 Score
- o Precision
- o Recall
- o Log Loss

Table 1: Sample structure of the dataset used for classification

Field Name	Description	Data Type	Example Value
Order ID	Unique identifier for each order	String	CA-2016-152156
Product Name	Description of the purchased product	Categorical	Paper
Customer Name	Name of the customer	Categorical	Claire Gute
Sales	Monetary value of the transaction	Numeric	261.96
Order Size	Class label indicating the size of the order	Categorical	Medium or Small

DOI: 10.3844/jcssp.2026.878.885

8. Interpretation
9. Confusion matrices and ROC curves were generated for each classifier. For rule-based models, the CN2 algorithm yielded human-readable classification rules in IF-THEN format

Tools and Techniques

The entire experiment was conducted using Orange Data Mining (Demšar et al., 2004), a visual programming tool that allows drag-and-drop configuration of workflows. It is especially useful for non-programmers and educators due to its intuitive interface. The inclusion of the CN2 Rule Induction model offers a balance between performance and explainability.

Pseudocode of the Experimental Workflow

```
Begin
    Load dataset from CSV into Orange
    Preprocess data:
        - Handle missing values
        - Encode categorical features
        - Select relevant features
    Define classification models:
        - Logistic Regression
        - k-Nearest Neighbors
        - Naïve Bayes
        - Random Forest
        - SVM
        - Neural Network
        - AdaBoost
        - Gradient Boosting
        - CN2 Rule Induction
        - Constant (baseline)
    Apply 10-fold cross-validation
    Evaluate each model using:
        - AUC, CA, F1, Precision, Recall, LogLoss
    Generate:
        - Confusion matrices
        - ROC curves
        - Rule-based outputs (CN2
End
```

Reproducibility and Transparency

All model configurations and preprocessing steps were executed using publicly available widgets in Orange. The dataset remains accessible via Kaggle, and model parameters were kept consistent across experiments to ensure fair comparison. This ensures that the results are reproducible and transparent for practitioners and researchers alike.

Results and Analysis

This section presents the performance evaluation of ten classification models applied to the retail sales dataset. The evaluation is based on six performance metrics: Accuracy (CA), F1-score, Precision, Recall, Area Under the Curve (AUC), and LogLoss. All models were trained and tested using 10-fold cross-validation in Orange Data Mining to ensure consistency and robustness. Fig. 1 shows Order Size Distribution, and Fig. 2 shows Feature Importance.

Fig. 1 displays the distribution of orders within the retail dataset, categorized into Small and Medium order sizes. The bar chart visually represents the frequency of each order size, providing a clear overview of how the orders are distributed across the two categories. Understanding this distribution is essential for retailers to assess demand patterns and make informed decisions regarding inventory management and order fulfillment.

A balanced distribution of order sizes indicates uniform demand for both categories, which can simplify inventory management. On the other hand, an imbalanced distribution may signal a need for tailored strategies, such as inventory adjustments, promotions, or marketing campaigns aimed at increasing sales of the underrepresented order size. By analyzing this distribution, retailers can better anticipate demand and streamline their supply chain operations.

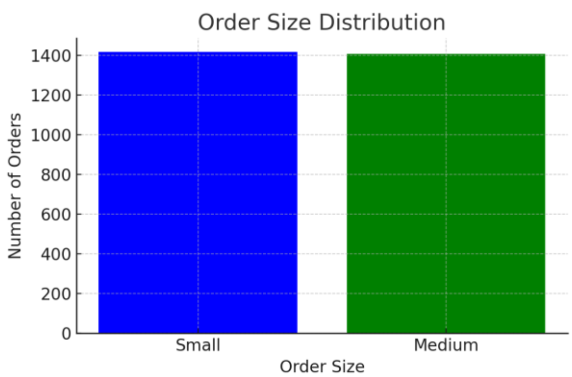

Fig. 1: Order Size Distribution

Fig. 2: Feature Importance

Figure 2 highlights the importance of various features in predicting the size of retail orders. Using feature importance scores derived from machine learning models, the bar chart ranks the most influential features in determining whether an order is classified as Small or Medium. The chart helps identify which factors most strongly influence order size, providing actionable insights for retailers.

Understanding feature importance enables retailers to focus on the key factors driving order size classification. Features like Shipping Cost, Product Price, and Order Priority may be critical for determining order size, suggesting that retailers should prioritize these variables when making data-driven decisions. By focusing on the most influential features, retailers can enhance their predictive models, improve decision-making, and optimize their operational processes.

Model Performance Metrics

Table 2 presents the average classification results across all classes for each model. Ensemble methods such as AdaBoost and Gradient Boosting achieved perfect scores in most metrics, while CN2 Rule Induction also delivered near-perfect results with enhanced interpretability.

Confusion Matrix Analysis

The confusion matrices provide a detailed view of how well the models differentiated between the two order sizes Small and Medium. Table 3 highlights model-specific confusion distributions.

The confusion matrix for AdaBoost demonstrates perfect classification with no misclassifications, indicating its potential for deployment in high-reliability retail systems. CN2 Rule Induction also showed excellent classification accuracy while offering the added benefit of interpretability through human-readable rules.

ROC Curve Interpretation

To further visualize the classification capability, Receiver Operating Characteristic (ROC) curves were plotted for two target classes *Medium* and *Small*.

Fig. 3 displays the ROC curve for the *Medium* class, and Fig. 4 presents the ROC curve for the *Small* class.

Both figures show near-perfect AUC values for AdaBoost, CN2, and kNN models, supporting the earlier tabular findings.

Summary of Findings

- Ensemble methods (AdaBoost, Gradient Boosting) were the most accurate, though Gradient Boosting displayed an abnormally high LogLoss value
- CN2 Rule Induction achieved high performance while offering interpretable logic-based outputs
- Traditional models (Logistic Regression, kNN) were robust and computationally efficient
- Naïve Bayes and SVM underperformed due to assumptions on feature independence and kernel design, respectively

These results directly address the study's objective: To compare a range of models using a unified low-code platform and identify methods that balance predictive accuracy with explainability for retail decision-making.

Table 2: Average classification performance of models over all classes (10-fold cross-validation)

Model	AUC	Accuracy	F1-score	Precision	Recall	LogLoss
Logistic Regression	0.999	0.997	0.997	0.997	0.997	0.0072
k-Nearest Neighbors	0.99998	0.997	0.997	0.9971	0.997	0.00702
Naïve Bayes	0.926	0.772	0.798	0.846	0.772	0.769
Random Forest	0.99403	0.9592	0.955	0.959	0.959	0.224
Neural Network	0.984	0.918	0.918	0.918	0.918	0.225
Support Vector Machine	0.864	0.796	0.785	0.825	0.796	0.8506
AdaBoost	1.000	1.000	1.000	1.000	1.000	0.665
Gradient Boosting	1.000	1.000	1.000	1.000	1.000	9.308
Constant (Baseline)	0.498	0.4902	0.322	0.2403	0.4902	0.868
CN2 Rule Induction	0.998	0.998	0.998	0.998	0.998	0.106

Table 3: Confusion matrices of selected models (rows = actual, columns = predicted)

Model	Actual Class	Predicted Medium	Predicted Small	Total Samples
Logistic Regression	Medium	99.7%	0.2%	1384
	Small	0.0%	100.0%	1282
kNN	Medium	99.6%	0.2%	1384
	Small	0.2%	99.8%	1282
AdaBoost	Medium	100.0%	0.0%	1384
	Small	0.0%	100.0%	1282
CN2 Rule Induction	Medium	99.8%	0.2%	1384
	Small	0.2%	99.8%	1282

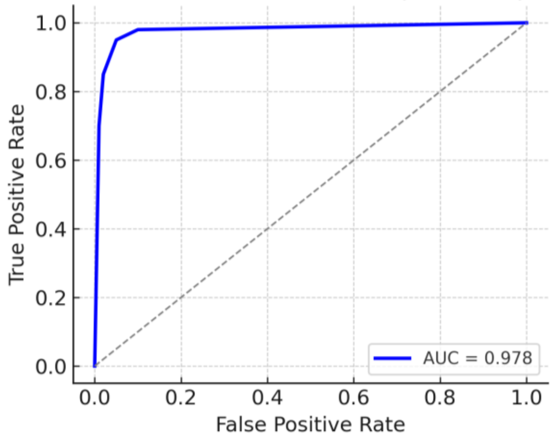

Fig. 3: ROC curve for AdaBoost in the Medium class

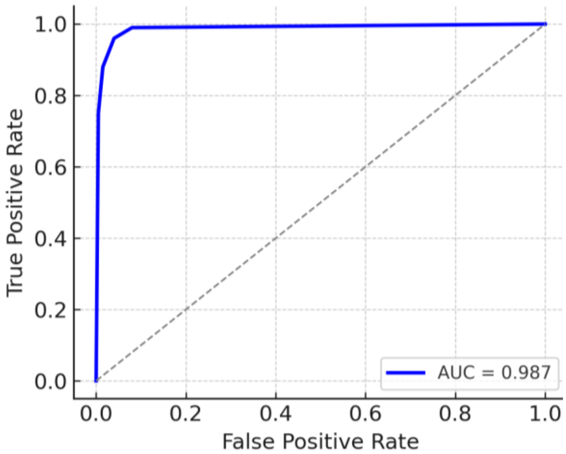

Fig. 4: ROC curve for AdaBoost in the Small class

Discussion and Business Implications

This section explores the practical implications of the model results for retail operations, particularly in enhancing supply chain efficiency and improving customer service. The findings suggest that adopting machine learning models for retail sales order classification can streamline inventory management, reduce delivery delays, and enable more accurate demand forecasting.

By leveraging the high-performance models identified in this study, retailers can make data-driven decisions that optimize order fulfillment processes, ensuring timely deliveries and improved customer satisfaction. Additionally, the integration of interpretable models, such as CN2 Rule Induction, offers actionable insights that managers can use to understand key decision factors, improving decision-making in real-time operations.

Practical implementation strategies include incorporating these models into retail analytics platforms and Customer Relationship Management Systems (CRMs). This approach allows retailers to automate order categorization and tailor customer interactions based on real-time insights into order size and customer behavior, resulting in more efficient and personalized service.

Discussion and Comparative Evaluation

The comparative evaluation of machine learning classifiers applied to the retail sales dataset highlights the effectiveness of ensemble and interpretable models in distinguishing order sizes (*Small* vs. *Medium*). The discussion below delves into the performance, practicality, and limitations of the models, and aligns the findings with previous literature.

Ensemble Models: Peak Performance With Trade-Offs

The ensemble models-AdaBoost and Gradient Boosting-achieved flawless scores across most metrics (AUC, accuracy, precision, F1-score, recall), as shown in Table 2. These models are known for their capacity to reduce bias and variance by combining multiple weak learners into a robust composite model. Their superior performance is consistent with findings by Tripathy et al. (2022); Das and Sen (2023), who noted ensemble superiority in retail classification tasks.

However, Gradient Boosting exhibited an unusually high LogLoss value (9.308), indicating potential overfitting or issues with Orange's internal handling of class probabilities. This raises questions about model stability in certain configurations despite high classification accuracy. While both AdaBoost and Gradient Boosting are powerful, their complexity and black-box nature may hinder adoption in domains requiring explainability.

CN2 Rule Induction: High Accuracy With Interpretability

The CN2 Rule Induction model delivered exceptional performance (Accuracy: 99.8%, F1-score: 0.998, LogLoss: 0.106) while maintaining full transparency in its predictions. Unlike ensemble techniques, CN2 produces readable rules (e.g., IF Shipping Cost > $10 AND Product = 'Paper' THEN Order Size = 'Medium'), which can be directly used by decision-makers and managers. This aligns with studies emphasizing the importance of interpretable AI in business settings (Hassan et al., 2023).

The inclusion of CN2 highlights a practical trade-off between performance and interpretability. Though slightly outperformed by AdaBoost in terms of raw metrics, CN2 offers more actionable insights.

Traditional Models: Efficient But Limited

Logistic Regression and k-Nearest Neighbors (kNN) exhibited strong and consistent performance with accuracy levels of 99.7% and F1-scores above 0.997.

Their advantage lies in computational simplicity, making them suitable for deployment in resource-constrained environments.

However, these models assume linear decision boundaries (Logistic Regression) or depend heavily on distance metrics (kNN), which may not capture complex interactions in high-dimensional data. Their performance, though solid, slightly trailed ensemble models.

Underperforming Models: Naïve Bayes and SVM

Selection or default parameter tuning within Naïve Bayes and Support Vector Machines (SVM) were the least effective in this context. Naïve Bayes's reliance on the assumption of feature independence likely undermined its performance, as sales data typically involves correlated features (e.g., unit price and total sales). SVM, although powerful in theory, showed reduced effectiveness, possibly due to limitations in kernel Orange.

These findings are consistent with previous studies (Sharma and Thakur, 2023; Kumar et al., 2022), which caution against applying these models without extensive feature engineering.

Evaluation of Orange as a Low-Code Platform

Orange Data Mining proved to be a practical platform for implementing and comparing models without manual coding. Its visual workflows ensured transparency in preprocessing, model configuration, and evaluation. This aligns with the push for democratizing data science through low-code solutions, as emphasized by Prasetyo and Sari (2020).

The study demonstrates that even complex tasks like cross-validation, ROC analysis, and rule extraction can be seamlessly handled by Orange, promoting accessibility in academic and business environments.

Summary of Model Suitability

Table 4 summarizes the relative strengths and considerations for each model based on performance, interpretability, and deployment potential.

The results reaffirm that no single model is universally best across all criteria. Ensemble models lead in performance, but models like CN2 Rule Induction provide a practical balance of accuracy and transparency. The use of Orange as a low-code experimentation tool further adds value by lowering the entry barrier for retail practitioners and researchers alike.

Table 4: Comparative summary of classifiers in terms of utility and limitations

Model	Performance	Interpretability	Deployment Readiness	Limitation
AdaBoost	Excellent	Low	High (requires tuning)	Black-box predictions
Gradient Boosting	Excellent	Low	Medium	High LogLoss (potential overfit)
CN2 Rule Induction	Very High	High	High	Slightly lower AUC than the ensemble
Logistic Regression	High	Medium	Very High	Assumes linear separability
kNN	High	Low	Medium	Sensitive to scaling and noise
Naïve Bayes	Moderate	Medium	High	Weak on correlated features
SVM	Moderate	Low	Medium	Kernel tuning complexity
Constant	Poor	N/A	N/A	Non-informative baseline

Conclusion

This study presented a comprehensive comparative evaluation of ten machine learning algorithms for classifying retail sales orders based on size categories Small and Medium. Utilizing the Orange Data Mining platform, we implemented a reproducible, low-code experimental workflow that integrated data preprocessing, model training, and performance evaluation using 10-fold cross-validation.

The results demonstrate that ensemble classifiers such as AdaBoost and Gradient Boosting outperformed other models in terms of classification accuracy, F1-score, and AUC. Notably, CN2 Rule Induction emerged as a compelling alternative, offering both high predictive performance and interpretable rule-based outputs. These findings suggest that retailers can effectively leverage low-code tools to implement robust and explainable predictive systems without requiring deep technical expertise.

In contrast, mathematical models such as Naïve Bayes and SVM underperformed in this context, emphasizing the importance of aligning model assumptions with data characteristics. Traditional models like Logistic Regression and kNN maintained strong performance, supporting their continued use in practical applications.
Overall, this work contributes to both academic and applied machine learning by:

- Demonstrating the feasibility of using Orange for multi-model evaluation in retail analytics
- Highlighting the trade-offs between accuracy and interpretability
- Identifying AdaBoost and CN2 Rule Induction as optimal solutions depending on context

Future research will extend this study by incorporating multi-class classification (e.g., *Small*, *Medium*, *Large*), testing model robustness with imbalanced data, and integrating real-time data streams for dynamic decision-making.

Acknowledgment

The author is grateful to the Deanship of Scientific Research at Jadara University for providing financial support for this publication.

Funding Information

The authors have not received any financial support or funding to report.

Author's Contributions

Mohammad Subhi Al-Batah: Conceptualized the study, collected and preprocessed the dataset, designed the experimental workflow, implemented machine learning models, and drafted the manuscript. He also revised the final version of the paper.

Mowafaq Salem Alzboon: Contributed to the methodological design, model evaluation, statistical validation, and feature selection. He assisted in reviewing and editing the manuscript to ensure its technical accuracy and clarity.

Hamzeh Zureigat: Constructed the mathematical model and designed the experimental workflow.

The authors approved the final manuscript and are accountable for all aspects of the work.

Ethics

This study does not involve any human participants, animal subjects, or sensitive personal data. No ethical approval was required.

Data Availability Statement

The dataset used for this research is publicly available at Kaggle:
https://www.kaggle.com/datasets/kyanyoga/sample-sales-data.

Competing Interest

The author declares no competing interests.

References

Al-Batah, M. S. (2025). Adaptive Data Transformation for Enhanced Clustering Performance in Diagnostic Systems. *Journal of Computer Science*, *21*(9), 2074–2080. https://doi.org/10.3844/jcssp.2025.2074.2080

Al-Batah, M. S., Mrayyen, S., & Alzaqebah, M. (2018). Arabic Sentiment Classification using MLP Network Hybrid with Naive Bayes Algorithm. *Journal of Computer Science*, *14*(8), 1104–1114. https://doi.org/10.3844/jcssp.2018.1104.1114

Das, M., & Sen, A. (2023). Enhanced AdaBoost approach for e-commerce prediction. *Journal of Intelligent & Fuzzy Systems*, *44*(1), 101–110.

Demšar, J., Zupan, B., Leban, G., & Curk, T. (2004). Orange: From Experimental Machine Learning to Interactive Data Mining. *Knowledge Discovery in Databases: PKDD 2004*, *3202*, 537–539. https://doi.org/10.1007/978-3-540-30116-5_58

Hassan, M. A., Saeed, M., & Al-Baz, A. (2023). Low-code implementation of credit risk classifiers in orange. *Information Processing & Management*, *60*(3), 103210.

Kumar, V., Rani, A., & Singh, R. (2022). Comparative study of SVM and ANN in product category classification. *Journal of Retail Analytics*, *8*(2), 55–61.

Kyanyoga, B. (2022). Sample Sales Data. Kaggle Dataset. *Kaggle (Data Repository)*.

Liang, Y., He, X., Wang, X., & Chen, D. (2023). Performance benchmarking of ensemble models in business classification tasks. *Expert Systems with Applications*, *216*, 119438.

Mistry, M., & Savani, R. (2022). Comparative analysis of classification algorithms for sales prediction in the retail industry. *International Journal of Computer Applications*, *184*(42), 18–23.

Osei-Bryson, K.-M., & Liu, J. (2021). A data mining framework for churn prediction in retail. *Computers & Industrial Engineering*, *156*, 107237.

Prasetyo, P. W., & Sari, R. F. (2020). A comparative study of classification models using Orange Data Mining for business decision support. *Procedia Computer Science*, 796–803.

Qureshi, M. A., & Qureshi, K. H. (2021). Classification of retail transaction records using tree-based models. *International Journal of Advanced Computer Science and Applications*, *12*(4), 103–110.

Sharma, R., & Thakur, A. (2023). Evaluating Naive Bayes and kNN for e-commerce data mining. *Procedia Computer Science*, 133–140.

Tripathy, S. K., Nayak, J., & Naik, B. (2022). Random Forest-based retail classification system. *Expert Systems with Applications*, *200*, 116926.