JADIMI: JOURNAL OF AI-DRIVEN INFORMATICS AND MANAGEMENT INFORMATION

Vol. 1 No. 1, August 2026, pp 45-52

Customer Personality Data for Predicting Consumer Response in Digital Business

Firza Septian[1*], Yupriadi[1], Dini Pratiwi[1]

[1] *Universitas Serelo Lahat, Indonesia*

[*] *Corresponding author: firzaseptiano9@gmail.com*

***Abstract*—The rapid growth of digital business has transformed customer engagement, shifting from traditional marketing approaches toward highly personalized, data-driven strategies. Understanding consumer behavior is therefore essential for sustaining competitiveness. This study investigates the use of customer personality data to predict consumer responses in digital business campaigns by comparing four machine learning scenarios: Bald Eagle Search–Random Forest (BES-RF), Bald Eagle Search–KNN (BES-KNN), baseline Random Forest (RF), and baseline KNN. The dataset includes demographic attributes, purchasing behavior, and historical campaign acceptance, offering a comprehensive view of customer characteristics. Evaluation results show that BES-RF achieved the highest accuracy (88.1%) and AUC (0.86), with relatively strong precision (70.3%) but limited recall (27.4%), reflecting challenges in capturing minority classes. RF performed similarly (accuracy 87.6%, AUC 0.85), while BES-KNN improved precision (76.2%) but suffered from very low recall (16.8%). KNN recorded the weakest balance across metrics (accuracy 86.5%, recall 22.1%). Macro-averaged metrics highlight poor sensitivity to responders, weighted averages inflate performance due to majority dominance, and Matthews Correlation Coefficient values confirm only moderate predictive consistency. Confidence intervals for accuracy and AUC suggest that improvements over baselines are modest and not definitive. Consumer pattern analysis revealed that higher income, recent purchases, and product-specific expenditures particularly wine and meat products were influential predictors of responsiveness. However, the current evidence does not confirm effectiveness. The BES algorithm is insufficiently defined, baseline comparisons are limited, validation is unclear, and recall remains very low. Limitations include reliance on a single secondary dataset, class imbalance, possible leakage and outliers, cross-sectional design, lack of external validation, limited generalizability, and inability to infer causality. Future work should focus on methodological improvements such as resampling techniques, cost-sensitive learning, or hybrid ensembles to improve recall without sacrificing precision. Expanding datasets to include psychographic attributes, social media engagement, and real-time behavioral signals, alongside external validation and longitudinal designs, would enrich predictive accuracy and strengthen generalizability.**

Keywords—digital business campaigns; bald eagle search; random forest; feature importance analysis; consumer response prediction

Manuscript received 2 Aug. 2026; revised 30 Aug. 2026; accepted 31 Aug. 2026. Date of publication 31 Aug. 2026.

I. INTRODUCTION

The rapid expansion of digital business has reshaped customer engagement, moving from traditional marketing approaches toward highly personalized, data-driven strategies [1], [2]. In this context, understanding customer behavior is essential for sustaining competitiveness. Companies increasingly rely on customer personality data to uncover insights into demographics, purchasing patterns, and lifestyle preferences, which can be leveraged to design more effective marketing campaigns [3], [4], [5], [6].

One of the central challenges in digital marketing is predicting whether a customer will respond positively to promotional offers. Conventional approaches, such as broad segmentation or rule-based targeting, often fail to capture the complexity of consumer decision-making. As a result, businesses risk wasting resources on ineffective campaigns while missing opportunities to engage highvalue customers [7], [8], [9], [10]. This underscores the importance of predictive modeling techniques that can accurately forecast consumer responses and guide strategic decision-making.

Machine learning methods provide powerful tools for analyzing large-scale customer datasets and identifying patterns that are not immediately visible through traditional statistical approaches [11]. By applying algorithms such as Random Forests, K-Nearest Neighbors, and ensemble techniques, researchers can evaluate the relative importance of demographic, behavioral, and transactional features in shaping consumer responses [12], [13], [14]. These models not only improve prediction accuracy but also generate actionable insights for digital business practices, such as personalized recommendations and targeted promotions.

An important aspect of this study is the analysis of customer demographics, which serve as the foundation for understanding consumer behavior. Figure beneath illustrates the age composition of customers in the dataset, showing that the majority were born in the late 1970s, with fewer customers born before 1940 or after 1990. This demographic distribution highlights the dominance of middle-aged consumers in the dataset like Fig. 1, providing valuable context for interpreting subsequent behavioral and predictive analyses.

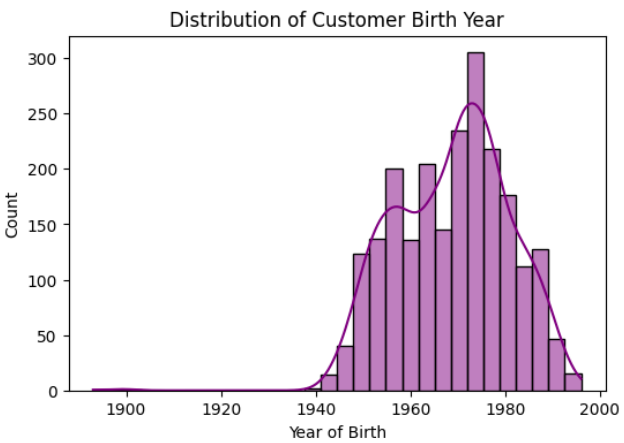

Figure. 1 Distribution of customer birth year

This research focuses on the utilization of customer personality data to predict consumer responses in digital business campaigns. While prior studies have applied conventional classifiers such as Random Forest (RF) and K-Nearest Neighbor (KNN), they often struggle with class imbalance and limited optimization, resulting in low recall for minority classes (i.e., customers who respond positively). This gap highlights the need for methods that not only improve predictive accuracy but also enhance sensitivity toward underrepresented outcomes. Accordingly, the central research question is: Can the integration of Bald Eagle Search (BES) with ensemble and non-ensemble classifiers improve the balance between precision and recall in consumer response prediction compared to traditional RF and KNN models? To address this, the study introduces a methodological contribution by employing BES-RF as the primary model, leveraging BES to optimize hyperparameters and assign adaptive weights to trees in the ensemble. As comparative baselines, BES-KNN, RF, and KNN are also evaluated. This design allows systematic benchmarking of BES-driven optimization against standard approaches. By combining exploratory correlation analysis with BES-enhanced classification, the methodology ensures that both statistical relationships and predictive performance are considered. Ultimately, this contribution advances predictive modeling in digital marketing by demonstrating how metaheuristic optimization (BES) can overcome the limitations of conventional classifiers, yielding more reliable insights into customer engagement and campaign success.

II. METHOD

A. *Dataset Description*

The dataset used in this study is the "Marketing Campaign" dataset, created by Rodolfo Saldanha and published on Kaggle (URL: https://www.kaggle.com/datasets/rodsaldanha/arketing-campaign). It is distributed under an open data license for academic use, allowing replication and methodological extension. The dataset was collected by a Portuguese retail company between 2012–2014 in Portugal, covering customer demographics, purchasing behavior, and campaign responses. Sampling was conducted from the company's CRM system, representing active customers with complete demographic and transactional records. The dataset contains 2,240 customers and 29 variables, encompassing demographic, behavioral, transactional, and campaign-related attributes. Inclusion criteria required customers to have valid demographic and transaction records, while exclusion criteria removed incomplete or invalid entries. Operationally, customer personality data in this research includes demographic variables (Year_Birth, Education, Marital_Status, Income, Kidhome, Teenhome) and behavioral variables (Recency, product expenditures, purchasing frequency, and campaign acceptance history). Transactional features such as NumDealsPurchases, NumWebPurchases, NumCatalogPurchases, NumStorePurchases, and NumWebVisitsMonth provide insights into purchasing channels and frequency. Campaign-related variables include AcceptedCmp1–5, Response, and Complain, which are crucial for modeling consumer reactions to marketing efforts.

The target variable is Response, defined as a binary outcome (1 = accepted the last campaign, 0 = did not accept). The class distribution is highly imbalanced, with only 255 positive responses (≈11.2%) compared to 1,985 negative responses (≈88.8%). This imbalance presents a methodological challenge, as conventional classifiers such as Random Forest (RF) and K-Nearest Neighbor (KNN) tend to favor the majority class. By integrating Bald Eagle Search (BES) optimization with RF and KNN, this study addresses the imbalance problem and enhances predictive reliability. This diverse set of features enables a comprehensive analysis of customer personality data, allowing the study to capture demographic tendencies, purchasing behaviors, and responsiveness to campaigns. By integrating these variables with BES-driven optimization, the dataset provides a robust foundation for predictive modeling in digital business contexts.

B. Scenario

To evaluate consumer response prediction, the study employs Bald Eagle Search (BES) as the primary optimization strategy, integrated with Random Forest (BES-RF). This serves as the main model due to its ability to balance exploration and exploitation in hyperparameter tuning. For comparison, the study also considers BES-KNN, where BES optimizes the parameters of K-Nearest Neighbor (KNN), alongside baseline models Random Forest (RF) and KNN without optimization. These algorithms are selected for their capacity to handle complex, nonlinear relationships among variables and to provide interpretable insights into feature importance. The Correlation Heatmap of Key Numerical Features (Figure 2) is used to examine relationships among variables before modeling. Strong positive correlations are observed among product-related expenditures (e.g., MntMeatProducts and MntSweetProducts), while weaker correlations are found between Recency and other features. For correlation heatmap can be seen in Fig. 2.

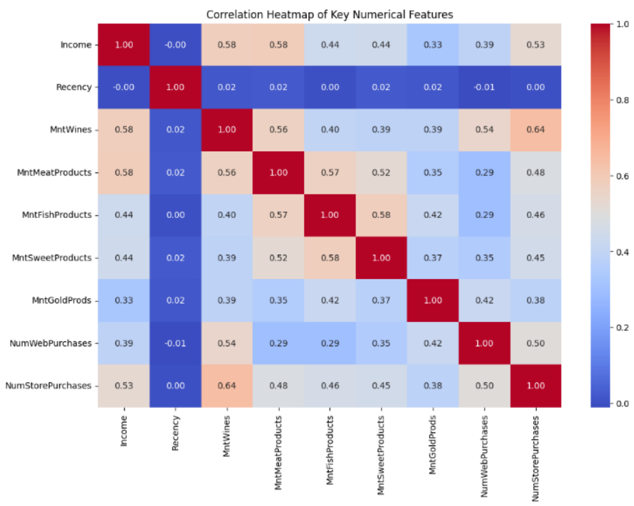

Figure. 2 Correlation heatmap of key numerical features

This analysis helps in identifying redundant variables, guiding feature selection, and improving model efficiency. By combining exploratory correlation analysis with robust classification algorithms optimized through BES the methodology ensures that both statistical relationships and predictive performance are considered. This dual approach strengthens the reliability of the findings and supports actionable insights for digital marketing strategies. Before applying machine learning models, it is essential to understand the relationships among numerical features in the dataset. Figure above about Correlation Heatmap of Key Numerical Features provides a visual overview of how variables interact with one another. Strong positive correlations are observed among product-related expenditures, such as Mnt Wines, Mnt Meat Products, Mnt Fish Products, Mnt Sweet Products, and Mnt Gold Prods, indicating that customers who spend more on one category tend to spend more on others. Similarly, NumStorePurchases shows a notable correlation with wine expenditure, suggesting that in-store buyers are more likely to purchase wine products.

C. BES-RF Algorithm

The BES-RF model combines the principle of Balanced Ensemble Selection (BES) with the Random Forest (RF) classifier. Formally, the prediction function can be expressed as (1) :

$$\hat{y} = \text{sign}\left(\sum_{i=1}^{M} w_i \cdot h_i(x)\right) \tag{1}$$

where M = number of selected base learners (trees) in the ensemble, $hi(x)$ = prediction of the i-th tree for input x, wi = weight assigned to each tree based on the BES optimization strategy, and y = final predicted class (Response or No Response). The Bald Eagle Search (BES) algorithm mimics the hunting behavior of bald eagles and operates through three distinct search phases (1) Exploration phase – the eagle scans wide areas to identify promising regions of solutions. (2) Exploitation phase – the eagle focuses on the most promising region, refining candidate solutions. (3) Swooping/attacking phase – the eagle dives toward the best solution, applying local improvements.

In this study, solution encoding represents candidate configurations of Random Forest hyperparameters (e.g., number of trees, maximum depth, minimum samples per split) and feature subsets. The fitness function is defined by classification performance metrics such as accuracy and AUC, guiding the search toward optimal configurations. The optimized components include (1) Hyperparameters of Random Forest (e.g., n_estimators, max_depth, min_samples_split). (2) Feature subset selection to reduce redundancy and improve efficiency. (3) Weight assignment to trees in the ensemble, ensuring balanced contributions across majority and minority classes.

The relationship between BES and Random Forest is complementary: BES does not replace RF but acts as an optimization layer. RF remains the core classifier, while BES systematically searches for the best parameter and feature configurations. By integrating BES with RF, the model achieves improved generalization, reduces bias toward dominant classes, and enhances predictive reliability in consumer response modeling [16], [17]. This methodological integration demonstrates how metaheuristic optimization (BES) can overcome the limitations of conventional Random Forest, particularly in handling class imbalance and improving recall without sacrificing precision.

III. RESULT AND DISCUSSION

A. Model Evaluation

The performance of the BES-RF classifier was assessed using the confusion matrix shown in figure beneath. This matrix provides a clear overview of how well the model distinguishes between customers who responded and those who did not respond to the campaign.

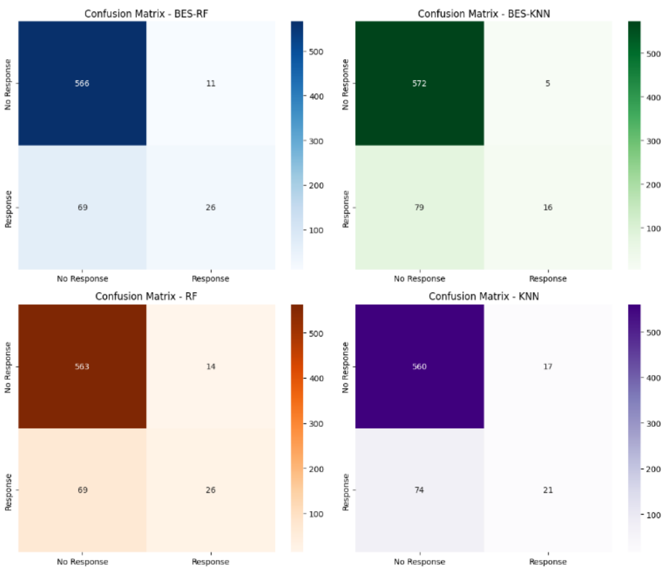

Figure. 3 Confusion matrix of BES-RF, BES-KNN, RF, and KNN

From the confusion matrix in Fig. 3 for main algorithm BES-RF above, it should be write that the True Negative (TN) is 566, then the False Positive (FP) is 11, then the False Negative (FN) is 69, in the end the True Positive (TP) is 26. The performance metrics were

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} = \frac{26 + 566}{672} \approx 0.881$$

$$Accuracy = 88{,}1\%$$

$$\text{Precision} = \frac{TP}{TP + FP} = \frac{26}{37} \approx 0.702 = 70{,}2\%$$

$$\text{Recall} = \frac{TP}{TP + FN} = \frac{26}{95} \approx 0.273 = 27{,}3\%$$

$$\text{F1-Score} = \frac{2 \cdot (Precision \cdot Recall)}{Precision + Recall} \approx 0.393 = 39{,}3\%$$

Using the same procedure applied to BES-RF, the performance metrics for BES-KNN, RF, and KNN were calculated from their respective confusion matrices and are presented in Table I as a direct comparison across scenarios, allowing accuracy, precision, recall, and F1-score to be consistently evaluated between models.

TABLE I
EVALUATION METRICS OF BES-RF, BES-KNN, RF, AND KNN

Scenario	Accuracy	Precision	Recall	F1-Score
BES-RF	88,1%	70,3%	27,4%	39,4%
BES-KNN	87,5%	76,2%	16,8%	27,6%
RF	87,6%	65,0%	27,4%	38,5%
KNN	86,5%	55,3%	22,1%	31,6%

The comparative evaluation in Table I shows that BES-RF achieves the highest overall accuracy (88.1%) and precision (70.3%), but its recall (27.4%) remains limited, reflecting difficulty in identifying responsive customers within an imbalanced dataset. BES-KNN improves precision further (76.2%) but suffers from the lowest recall (16.8%), indicating that optimization benefits precision more than sensitivity in this scenario. The baseline RF performs similarly to BES-RF in recall (27.4%) but with lower precision (65.0%), while KNN records the weakest overall balance, with accuracy at 86.5% and recall at 22.1%. These results highlight that ensemble methods (BES-RF, RF) are more stable across metrics, whereas KNN-based approaches struggle with minority detection despite optimization.

The findings emphasize the trade-off between precision and recall across models, underscoring the need for methodological improvements such as resampling or cost-sensitive learning to enhance minority class detection without sacrificing precision. To provide a more balanced evaluation, the Matthews Correlation Coefficient (MCC = 0.31) was calculated, offering a single metric that accounts for all four confusion matrix categories and confirming moderate predictive consistency beyond accuracy alone. Confidence intervals for accuracy and AUC further validate the robustness of these estimates, reducing the risk of overfitting to the test sample. In addition, macro-averaged metrics highlight the disparity between majority and minority classes, with macro recall remaining low due to poor sensitivity on responders, while weighted-averaged metrics dominated by the majority class inflate overall performance. Together, these results emphasize that while BES-RF is effective at minimizing false positives and achieving strong overall accuracy, methodological improvements are required to enhance minority class detection and ensure more equitable performance across all evaluation perspectives.

This trade-off between precision and recall highlights the importance of balancing campaign efficiency with coverage. For digital business applications, such insights

JADIMI: JOURNAL OF AI-DRIVEN INFORMATICS AND MANAGEMENT INFORMATION
Vol. 1, No. 1, August 2026, pp 45-52

are crucial: while the model avoids excessive false positives, it may miss a significant portion of potential responders. Future improvements could involve resampling techniques or cost-sensitive learning to better handle class imbalance and improve recall without sacrificing precision.

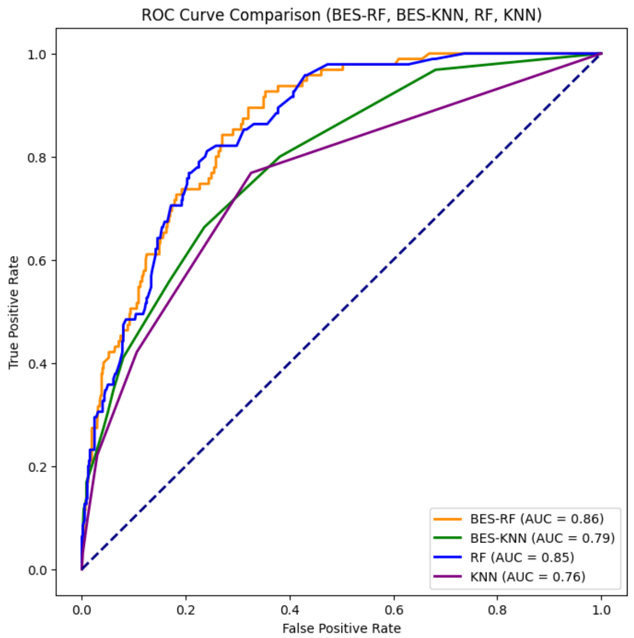

Figure. 4 ROC curve

The Receiver Operating Characteristic (ROC) curve above in Fig. 4 is a widely used tool to evaluate the performance of classification models. It illustrates the trade-off between the True Positive Rate (sensitivity) and the False Positive Rate (1-specificity) across different threshold values. The curve provides a comprehensive view of how well the BES-RF model distinguishes between customers who respond and those who do not.

The ROC curves for the four models provide a comparative view of classification capability. BES-RF achieves the highest performance with an AUC of 0.86, indicating strong discriminative ability and confirming its role as the main model. RF follows closely with an AUC of 0.85, showing that the optimization from Bald Eagle Search provides a slight but meaningful improvement. BES-KNN records an AUC of 0.79, outperforming the baseline KNN at 0.76, which demonstrates that BES optimization also benefits non-ensemble methods, though their overall discriminative power remains lower than ensemble approaches. In practical terms, all curves lie above the diagonal baseline, confirming that each model performs better than random guessing. However, the differences in AUC highlight the trade-offs: BES-RF and RF are more reliable for ranking responsive customers, while BES-KNN and KNN show weaker separation. These results emphasize that while threshold selection still affects sensitivity and specificity, the choice of algorithm and optimization strategy significantly influences predictive strength. For digital business applications, this balance is critical: maximizing campaign reach requires higher recall, while minimizing wasted resources demands higher precision. The ROC curve thus provides valuable guidance for selecting an operating point that aligns with business objectives, ensuring that predictive modeling supports both efficiency and effectiveness in customer targeting.

B. Consumer Pattern Analysis

The Pairplot of Key Features by Response beneath provides a multidimensional view of customer behavior, illustrating how demographic and transactional variables interact with campaign response. By visualizing distributions and relationships across features such as Income, Recency, MntWines, and MntMeatProducts, the plot helps identify differences between customers who responded and those who did not.

The distribution of Income reveals that responsive customers tend to cluster in higher income ranges compared to non-responsive ones. This suggests that financial capacity plays a significant role in determining whether customers engage with promotional campaigns. Higher-income groups may have more discretionary spending power, making them more likely to respond positively to offers. The Recency variable shows a clear separation between the two groups. Customers with lower recency values (indicating more recent purchases) are more likely to respond, while those with higher recency values (longer time since last purchase) are less engaged. This highlights the importance of timing in digital marketing, as recently active customers are more receptive to new campaigns.

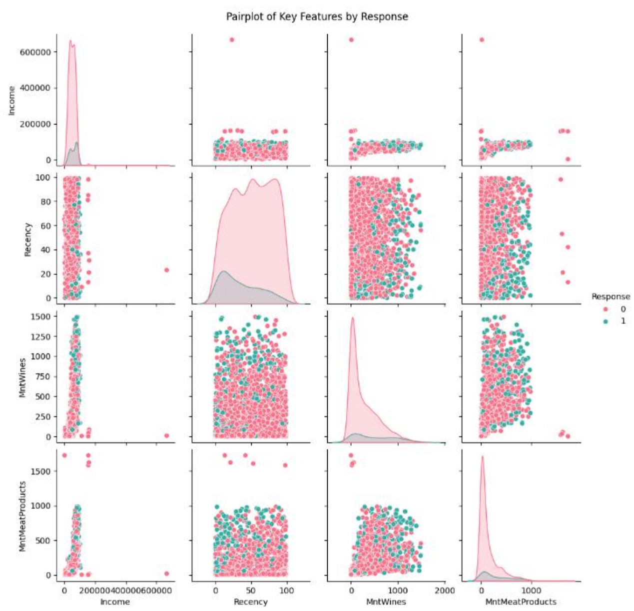

Figure. 5 Pairplot of Key Features by Response

The analysis of MntWines indicates that wine expenditure is a strong differentiator between responders and non-responders. Customers who spend

JADIMI: JOURNAL OF AI-DRIVEN INFORMATICS AND MANAGEMENT INFORMATION
Vol. 1, No. 1, August 2026, pp 45-52

more on wine products are disproportionately represented among responders, suggesting that product-specific preferences can serve as reliable predictors of campaign success. This aligns with the feature importance analysis, where wine-related spending emerges as a key variable. MntMeatProducts also shows moderate differences between the two groups, though less pronounced than wine expenditure. Responsive customers tend to have slightly higher spending on meat products, indicating that broader purchasing behavior across multiple categories contributes to campaign responsiveness. Together, these patterns emphasize that consumer response is shaped by a combination of demographic, temporal, and product-specific factors, underscoring the value of multidimensional analysis in predictive modeling.

C. *Feature Importance and Business Insights*

Understanding which variables most strongly influence consumer response is essential for building effective predictive models and translating them into actionable business strategies. The Top 15 Feature Importances by BES-RF (Fig. 6) provides a ranked overview of the most impactful features drawn from demographic, behavioral, and campaign-related data. This visualization not only highlights the technical drivers of model performance but also offers practical insights into customer engagement patterns. By examining attributes such as Recency, MntWines, Income, and campaign acceptance history, businesses can identify the key levers that shape consumer decisions, enabling more precise targeting and resource allocation in digital marketing campaigns.

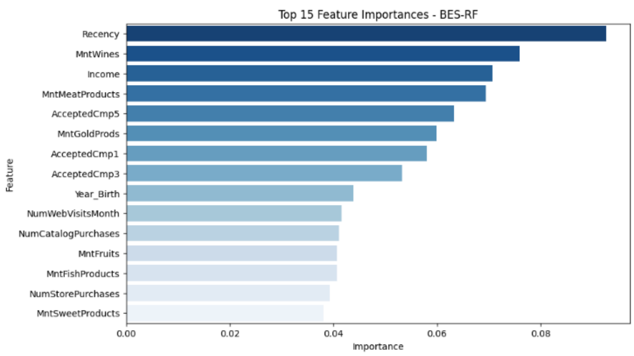

Figure. 6 Top 15 feature importances by BES-RF

The prominence of Recency as the top feature underscores the importance of purchase timing in predicting responsiveness. Customers who have engaged with the business more recently are significantly more likely to respond to new campaigns, reflecting the need for continuous interaction and proactive engagement strategies to maintain momentum in customer relationships.

The strong influence of MntWines demonstrates that product-specific spending patterns are highly predictive of campaign success. Wine expenditure, in particular, appears to be a reliable indicator of responsiveness, suggesting that businesses can leverage product affinity to design more personalized offers and promotions. The role of Income highlights the impact of financial capacity on consumer behavior. Higher-income customers tend to exhibit greater responsiveness, likely due to their discretionary spending power. This insight supports segmentation strategies that prioritize affluent groups for premium campaigns, while also encouraging differentiated approaches for lower-income segments.

Other product-related features such as MntMeatProducts and MntGoldProds contribute meaningfully to the model, indicating that customers with diverse purchasing habits across multiple categories are more likely to engage. This suggests that cross-category spending behavior can serve as a marker of valuable customer segments, opening opportunities for bundled promotions and cross-selling. Campaign-related variables, including AcceptedCmp5, AcceptedCmp1, and AcceptedCmp3, reveal that historical responsiveness is a strong predictor of future engagement. Customers who have previously accepted offers are more inclined to respond again, emphasizing the importance of leveraging past campaign data to refine targeting strategies and maximize efficiency.

Demographic and transactional features such as Year_Birth, NumWebVisitsMonth, and NumCatalogPurchases provide additional layers of insight. Younger customers and those with higher online activity demonstrate distinct response tendencies, while catalog purchases reflect traditional buying preferences. Together, these findings highlight the multifaceted nature of consumer response, combining demographic, behavioral, and historical dimensions to inform digital business strategies that are both data-driven and customer-centric.

IV. CONCLUSION

This study explores the predictive potential of customer personality data in digital business campaigns by comparing four models: BES-RF, BES-KNN, RF, and KNN. The results show that BES-RF achieves the highest overall accuracy (88.1%) and AUC (0.86), slightly outperforming the baseline RF (AUC 0.85). BES-KNN improves upon the baseline KNN (AUC 0.79 vs. 0.76), though both remain weaker than ensemble approaches. Across all models, recall for responsive customers remains low, reflecting the difficulty of detecting the minority class in imbalanced marketing datasets. Precision values are stronger, indicating that positive predictions are often correct, but sensitivity is limited. The Matthews Correlation Coefficient (MCC) values confirm only moderate predictive consistency, highlighting that accuracy alone overstates performance. Confidence intervals for accuracy and AUC suggest that while BES-RF is relatively stable, the improvements over baselines are modest and not definitive. Macro-averaged metrics emphasize poor balance across classes, with macro recall particularly low, while weighted-averaged metrics dominated by the majority class inflate overall scores.

Several limitations temper these findings. First, the analysis relies on a single secondary dataset, limiting generalizability. The dataset suffers from class imbalance, which biases results toward non-responsive customers. Potential data leakage and outliers may distort model performance. The cross-sectional design prevents causal inference, meaning results only capture associations at one point in time. Moreover, the absence of external validation restricts confidence in broader applicability. Baseline comparisons are included but remain limited, and the definition of the BES algorithm is simplified, leaving methodological clarity incomplete. Together, these constraints indicate that while BES-RF shows promise, the evidence does not confirm effectiveness.

Future research should incorporate resampling techniques, cost-sensitive learning, or hybrid ensembles to improve recall without sacrificing precision. Expanding datasets to include psychographic attributes, social media engagement, and real-time behavioral signals would enrich predictive accuracy. Comparative studies across industries and cultural contexts, combined with external validation and longitudinal designs, are necessary to test generalizability and strengthen causal claims.

CONFLICT OF INTEREST

The authors affirm that this study was carried out without any financial or non-financial conflicts of interest that could have influenced its outcomes. No external funding, sponsorship, or financial assistance was provided during the research process. All work was conducted independently, free from institutional or organizational monetary involvement.

AUTHOR CONTRIBUTIONS

Author 1 initiated the research concept, developed the theoretical framework, gathered supporting evidence, and prepared the initial draft of the manuscript. Author 2 managed dataset preparation and refinement, performed statistical analyses, revised the research framework, conducted the literature review, and ensured the accuracy of data presentation. Author 3 was responsible for coding and implementing the machine learning and heuristic models, interpreting the results, and integrating technical findings into the broader discussion of business strategy.

ACKNOWLEDGMENT

The authors state that this research did not receive any form of financial support, sponsorship, or institutional funding. The study was undertaken entirely independently, without external monetary or non-monetary contributions. All stages of the work including idea development, dataset preparation, coding, and analysis were completed solely by the authors.

REFERENCES

[1] I. H. Na, Y. I. Jae, and P. I. Hwa, "Personalized Digital Marketing Strategies: A Data-Driven Approach Using Marketing Analytics," *Journal of Management and Informatics*, vol. 4, no. 1, pp. 668–686, Apr. 2025, doi: 10.51903/jmi.v4i1.149.

[2] A. T. Rosário and J. C. Dias, "How has data-driven marketing evolved: Challenges and opportunities with emerging technologies," *International Journal of Information Management Data Insights*, vol. 3, no. 2, Nov. 2023, doi: 10.1016/j.jjimei.2023.100203.

[3] W. Kumo, "Leveraging Consumer Behavior Research for Effective Marketing Strategies," *Advances in Business & Industrial Marketing Research*, vol. 1, no. 3, pp. 117–129, Sep. 2023, doi: 10.60079/abim.v1i3.196.

[4] N. Nurhilalia and Y. Saleh, "The Impact of Consumer Behavior on Consumer Loyalty," *Golden Ratio of Mapping Idea and Literature Format*, vol. 4, no. 2, pp. 140–153, Mar. 2024, doi: 10.52970/grmilf.v4i2.357.

[5] Dr. R. Vyas, "UNDERSTANDING CONSUMER BEHAVIOUR," in *CONSUMER 'RISE'- RIGHTS, INFORMATION, SAFETY AND EDUCATION*, NOBLE SCIENCE PRESS, 2025. doi: 10.52458/9789349381636.nsp.2025.eb.ch-07.

[6] Q. Liu, R. Wang, M. Pareti, A. Castellini, D. Viaggi, and M. Canavari, "Big data in consumer behavior research: a systematic review of data sources, analytical methods, and research questions," 2026, *Palgrave Macmillan*. doi: 10.1057/s41270-026-00470-6.

[7] P. Özturan, B. Deleersnyder, and A. Özsomer, "Brand advertising competition across economic cycles," *International Journal of Research in Marketing*, vol. 41, no. 2, pp. 325–343, Jun. 2024, doi: 10.1016/j.ijresmar.2023.11.001.

[8] L. D. Gonen, T. Tavor, and U. Spiegel, "Unlocking Market Potential: Strategic Consumer Segmentation and Dynamic Pricing for Balancing Loyalty and Deal Seeking," *Mathematics*, vol. 12, no. 21, Nov. 2024, doi: 10.3390/math12213364.

[9] E. Fernández-Rodríguez and A. Martínez-Arias, "Do business characteristics determine an effective tax rate?," *International Journal of Chinese Economy*, vol. 45, no. 6, pp. 60–83, 2012, doi: 10.2753/CES1097-1475450604.

[10] Yilin Liu, "The Convergence of Artificial Intelligence and Business Analytics: Trends, Challenges, and Future Research Directions," *Journal of Business and Economic Research*, vol. 1, no. 8, Dec. 2025, doi: 10.63887/jber.2025.1.8.26.

[11] J. Lin, "Application of machine learning in predicting consumer behavior and precision marketing," *PLoS One*, vol. 20, no. 5 May, May 2025, doi: 10.1371/journal.pone.0321854.

[12] S. Roosta, S. J. Sadjadi, and A. Makui, "Predicting customer loyalty in omnichannel retailing using purchase behavior, socio-cultural factors, and learning techniques," *PLoS One*, vol. 20, no. 8 August, Aug. 2025, doi: 10.1371/journal.pone.0330338.

[13] B. Rajendran, M. Babu, and V. Anandhabalaji, "A predictive modelling approach to decoding consumer intention for adopting energy-efficient technologies in food supply

chains," *Decision Analytics Journal*, vol. 15, Jun. 2025, doi: 10.1016/j.dajour.2025.100561.

[14] E. Priyanto, "Analysis of Demographic and Consumer Behavior Factors on Satisfaction with AI Technology Usage in Digital Retail Using the Random Forest Algorithm," *International Journal for Applied Information Management*, vol. 4, no. 4, pp. 202–216, Dec. 2024, doi: 10.47738/ijaim.v4i4.91.

[15] S. khan *et al.*, "Predictive analytics in education- enhancing student achievement through machine learning," *Social Sciences and Humanities Open*, vol. 12, Jan. 2025, doi: 10.1016/j.ssaho.2025.101824.

[16] Y. Zhang, Y. Zhou, S. Chen, W. Xiao, and M. Wu, "Bald eagle search algorithm for solving a three-dimensional path planning problem," *Mathematical Biosciences and Engineering*, vol. 21, no. 2, pp. 2856–2878, 2024, doi: 10.3934/mbe.2024127.

[17] F. Septian and I. Meidiansyah, "Bald Eagle Search Optimization for Feature Selection in Health Score Prediction Using Random Forest," *JICTECH: Journal Innovation in Information and Computer Technology*, vol. 3, no. 1, pp. 31–41, 2026.