

Research Article

Machine Learning-Based Vehicle Renewal Prediction: A Hybrid Approach for Customer Retention in Premium Automotive Markets

Selçuk Bayracı ^{1*}, Garen Bozoğlanoğlu ², Turgay Tugay Bilgin ³

¹ Borusan Otomotiv R&D Center, 0000-0003-4831-4802, selcuk.bayraci@borusanotomotiv.com

² Borusan Otomotiv R&D Center, 0009-0001-4977-6972, garen.bozoglanoglu@borusanotomotiv.com

³ Bursa Technical University, 0000-0002-9245-5728, turgay.bilgin@btu.edu.tr

* Correspondence: selcuk.bayraci@borusanotomotiv.com ; +90 532 2020975

Received: 25 May 2025

Revised: 14 October 2025

2nd Revised: 22 October 2025

3rd Revised: 02 December 2025

Accepted: 03 December 2025

Published: 07 December 2025

This is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Reference: Bayracı, S., Bozoğlanoğlu, G., & Bilgin, T. T. (2025). Machine learning-based vehicle renewal prediction: A hybrid approach for customer retention in premium automotive markets. *The European Journal of Research and Development*, 5(1), 355–377.

Abstract

Customer retention and vehicle renewal prediction remain critical challenges in premium automotive markets. This study presents a comprehensive data-driven framework for predicting BMW customer renewal probability using historical transactional and behavioral data from Borusan Otomotiv's enterprise systems. We developed a hybrid machine learning model that integrates Random Forest feature selection with Binary Logistic Regression to achieve interpretability while maintaining predictive accuracy. The model leverages customer demographics, service engagement metrics, and ownership patterns to generate individual-level renewal probability scores.

Evaluated on 1,211 holdout observations through temporal validation, the model achieved 77% overall accuracy and an AUC-ROC of 0.80, demonstrating strong discriminatory power in distinguishing between renewal and non-renewal customers. Model outputs are transformed into

five operational risk grades (G1-G5) and seamlessly integrated into Salesforce CRM, enabling proactive customer relationship management and targeted retention strategies.

Key empirical findings indicate that service expenditure patterns, time since last purchase, and multi-vehicle ownership significantly influence renewal likelihood. The framework bridges predictive analytics with operational deployment through automated data pipelines and continuous model monitoring, representing a practical approach to data-driven customer retention in the automotive sector.

Keywords: Customer retention prediction, Vehicle renewal forecasting, Automotive analytics, Machine learning, Logistic regression, Predictive modeling, Random Forest.

1. Introduction

Customer retention and vehicle renewal represent key performance drivers in the premium automotive sector. As competition intensifies and consumer loyalty becomes increasingly dynamic, understanding and predicting customer renewal behavior has become a strategic imperative for automotive companies. In particular, luxury brands such as BMW face the dual challenge of sustaining customer engagement across product lifecycles and optimizing marketing investments toward customers most likely to repurchase.

In this context, predictive analytics offers a powerful means to enhance customer lifecycle management by transforming historical behavioral and transactional data into actionable insights. Machine Learning (ML) models, when effectively designed and deployed, can uncover subtle patterns in service, purchase, and engagement data that are indicative of future renewal intentions. Such models enable organizations to identify at-risk customers, personalize retention campaigns, and prioritize sales leads with the highest likelihood of conversion.

The present study introduces a data-driven framework for estimating the vehicle renewal probability of existing BMW customers, termed the Personalized Campaign Offerings (PCO) Prediction Model. The model leverages structured data from BMW's enterprise systems including customer account information, vehicle ownership records, and service transactions to forecast the probability of a subsequent vehicle purchase. By combining behavioral and transactional indicators with statistical and machine learning methods, the proposed model provides a transparent and interpretable mechanism for renewal prediction at the individual customer level.

Unlike purely heuristic or rule-based retention tools, this approach integrates variable selection via ensemble learning (Random Forests) with a parametric, interpretable predictive structure (Binary Logistic Regression). This hybrid design balances predictive accuracy with operational interpretability, enabling business users to understand the key drivers of renewal while maintaining statistical rigor.

The contribution of this study is threefold. First, it presents a comprehensive modeling pipeline from data curation and leakage mitigation to temporal validation and operational deployment tailored to the automotive retail context. Second, it empirically demonstrates how service engagement metrics and ownership behaviors influence renewal probabilities, offering insights for data-driven customer relationship management. Third, it proposes a practical implementation framework through which the model's outputs are transformed into operational risk grades, thereby supporting proactive sales and marketing strategies.

The remainder of this paper is organized as follows. Section 2 reviews related work on predictive modeling for customer retention and renewal in the automotive and retail sectors. Section 3 outlines the methodological framework, detailing the data sources, model specification, and validation procedures. Section 4 presents the empirical results and performance evaluation. Section 5 discusses the managerial implications of the findings, and Section 6 concludes with recommendations for future research and deployment perspectives.

2. Literature Review: Predictive Modeling in the Automotive Sector

2.1 Introduction and Scope of Review

Accurate demand and purchase prediction models are critical for the efficiency of the automotive supply chain, minimizing inventory costs, and maximizing sales effectiveness. This systematic literature review synthesizes recent empirical work on predictive modeling within the automotive sector, focusing on the techniques, data inputs, and specific predictive objectives. The review is structured to identify current research trends and highlight the methodological gap addressed by the present study: the firm-level prediction of customer renewal likelihood based on historical customer data.

The reviewed literature primarily falls into three thematic categories: (1) Macro-level aggregate sales forecasting, (2) Micro-level customer purchase classification and intent prediction, and (3) Advanced modeling approaches for specific targets.

2.2 Macro-level Aggregate Sales Forecasting

This stream of research focuses on predicting total market or regional sales volume, often relying on macroeconomic and external indicators. Kaya and Yildirim (2020) proposed an 8-layer Deep Neural Network (DNN) for automobile sales prediction in Turkey, successfully utilizing macroeconomic variables such as exchange rate, GDP, consumer confidence index, and consumer price index [4]. This highlights the effectiveness of Deep Learning architectures in capturing complex, non-linear dependencies among aggregated time-series data.

Wachter et al. (2019) demonstrated the significant predictive power of customer pre-purchase online search data, combining keyword-specific customer online search data with traditional economic variables [10]. Their approach resulted in quantifiable reduction in out-of-sample error compared to models relying solely on economic indicators, underscoring the value of real-time digital behavior as an exogenous predictor.

2.3 Micro-level Customer Purchase Classification and Intent

This body of work focuses on individual customer behavior, aiming to classify potential purchasers or predict their intent. Wah and Ismail (2011) conducted a comparative study on predicting customer car purchase intent following a car booking, testing Logistic Regression, Decision Tree, and Neural Network models [11]. While all three achieved high validation accuracy rates (above 91%), Logistic Regression demonstrated the highest sensitivity, suggesting its practical utility in identifying positive purchase intent cases. More recent applications confirm the continued relevance of Logistic Regression for customer behavior prediction, with studies demonstrating its effectiveness when enhanced with regularization techniques (L1 and L2) to improve precision, recall, and overall model accuracy [15].

Yin et al. (2018) addressed the classification of automobile purchases using a Deep Learning method enhanced by network information, demonstrating that social and interaction networks can provide powerful latent features that traditional customer records miss [12]. Das Mou et al. (2021) provided a comprehensive review of supervised machine learning algorithms for predicting initial car purchase decisions, confirming the versatility of classification techniques like Naive Bayes and k-Nearest Neighbor [2].

Recent advances in automotive customer retention demonstrate the growing sophistication of predictive modeling approaches. A comprehensive comparative study

by Sharma et al. (2025) evaluated traditional machine learning models against ensemble methods (XGBoost, Random Forest) for churn prediction, confirming that ensemble models achieve near-perfect results across accuracy, precision, recall, and F1 scores [8]. This validates hybrid approaches combining multiple modeling techniques.

Specifically focusing on automotive after-sales service, Dinasti et al. (2025) integrated big data analytics with ensemble machine learning for vehicle service churn prediction, emphasizing the critical role of service engagement in retention [3]. Their findings align with earlier work demonstrating that service visit patterns strongly predict renewal likelihood. Liu et al. (2024) advanced this further by demonstrating extreme gradient boosting with Bayesian optimization for profit-driven churn prediction, achieving superior results over traditional approaches [6]. Similarly, recent work on AI-driven predictive maintenance in the automotive sector demonstrates that timely service appointment scheduling and maintenance prediction can increase customer service visit rates by 30-37%, highlighting the bidirectional relationship between service engagement and customer retention [14].

A systematic review by Manzoor et al. (2024) covering 240 peer-reviewed studies published between 2020-2024 identified increasing preference for ensemble models, profit-driven frameworks, and hybrid architectures in churn prediction across telecommunications, retail, banking, and automotive sectors [7]. The review emphasized class imbalance, concept drift, and model interpretability as persistent challenges requiring attention. Thangeda et al. (2024) specifically employed neural network-based predictive decision models for automotive customer retention, demonstrating the sector-specific application of advanced techniques [9].

A comprehensive survey covering churn analysis studies across multiple sectors emphasizes that customer churn prediction has evolved from simple statistical methods to sophisticated machine learning approaches, with particular attention to model interpretability and feature selection challenges [13].

This growing body of evidence supports the effectiveness of machine learning in automotive customer retention while highlighting the importance of interpretability and operational integration factors central to the current study.

2.4 Advanced Modeling and Scope Refinement

A smaller subset of research focuses on advanced hybrid models and more granular, quantitative predictions. Kim et al. (2023) tackled the challenge of firm-level demand

forecasting, establishing a machine learning-assisted hybrid-input model integrating both firm-specific (endogenous) and market-wide (exogenous) data [5]. This approach represents a more complex, business-relevant scope by focusing on a single manufacturer or brand rather than aggregate market prediction.

Chen et al. (2024) addressed quantitative prediction of car purchase amount using a BP Neural Network optimized by Genetic Algorithm, demonstrating that optimization techniques can significantly enhance neural network performance for regression tasks in the automotive domain [1].

2.5 Synthesis and Research Gap

The systematic review reveals several key findings: (1) Dominance of classification models with majority of micro-level studies utilizing LR, DT, NN, and DL to predict binary outcomes; (2) Diverse data integration where predictive accuracy is significantly enhanced by incorporating non-traditional data sources, particularly macroeconomic indicators and digital footprint data; (3) Methodological progression from classical models to more complex Deep Learning and algorithmically optimized hybrid approaches.

The current study focuses on renewal prediction, a specialized micro-level classification task distinct from initial purchase intent or general market forecasting. While Kim et al. (2023) address firm-level demand and Wah and Ismail (2011) address post-booking intent, there is a noted gap regarding a dedicated model for predicting the specific event of a current customer renewing their lease or replacing their vehicle using their own historical transaction and service engagement data [5] [11]. Our approach, utilizing a Logistic Regression model on a temporally partitioned, proprietary dataset, fills this gap by providing an interpretable, operational prediction score for an existing customer's repurchase cycle.

3. Methodology

This section describes the end-to-end methodological framework employed for developing and deploying the BMW Customer Renewal Probability Prediction Model. The system architecture integrates data engineering, security governance, machine learning development, and operational deployment within a unified cloud-based infrastructure. The workflow comprises seven integrated stages as illustrated in Figure 1.

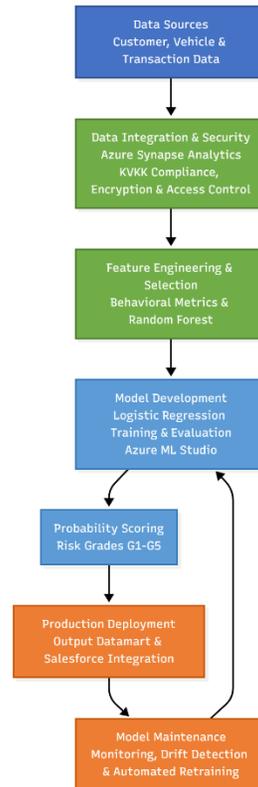


Figure 1: System Architecture Flowchart

The process begins with Data Sources encompassing three enterprise systems: Customer Master Data containing demographic and account information, Vehicle Ownership Data tracking registration and purchase history, and Sales and Service Transaction logs recording all customer interactions.

Data Integration & Security is performed through Azure Synapse Analytics, which consolidates heterogeneous sources through ETL processing while implementing KVKK Compliance protocols for personal data protection, encryption at rest and in transit, and role-based access control. This governance framework ensures regulatory compliance while maintaining data utility for predictive modeling.

Feature Engineering & Selection transforms raw transactional data into predictive features through temporal aggregations and behavioral metric calculations. Random Forest importance analysis identifies 15 influential predictors exceeding the 0.02 threshold, including time since last purchase, service engagement metrics, and customer portfolio characteristics, balancing predictive power with interpretability.

Model Development occurs within Azure ML Studio using Binary Logistic Regression with L2 regularization. The model achieves 72% accuracy, 70% precision and recall, and 0.80 AUC-ROC on the holdout test set, demonstrating strong discriminatory capability in distinguishing renewal from non-renewal customers.

Probability Scoring transforms continuous predictions into five operational risk grades (G1-G5) through optimal binning, enabling sales teams to prioritize high-probability customers (G1-G2) for retention campaigns while implementing specialized strategies for lower-probability segments.

Production Deployment stores validated scores in the Output Datamart and integrates them into Salesforce Customer Cards through automated API synchronization, enabling sales teams to access renewal probabilities within existing CRM workflows.

Model Maintenance implements continuous monitoring dashboards tracking prediction distributions, feature stability, and business outcomes. Automated retraining pipelines refresh the model monthly with recent data, ensuring sustained predictive accuracy as customer behaviors evolve.

3.1 Data Architecture and Sources

All data utilized in this study originate from Borusan Otomotiv's enterprise systems, stored and managed within the company's Azure data environment. Three primary systems were integrated:

- **Azure SQL Database:** Serves as the foundational data source containing customer master data, vehicle ownership records, and transactional histories.
- **Azure Synapse Analytics:** Acts as the central analytical warehouse and model development environment, facilitating large-scale data transformation, feature engineering, and temporal aggregation.
- **Azure Machine Learning Studio:** Provides the experimentation and model management environment for training, validation, and deployment of the predictive models.

The analysis is based on a dataset of 6,027 BMW 5 Series customers observed over a three-year period (October 2021 to October 2024), all of whom had purchased at least one BMW 5 Series vehicle during this timeframe. The empirical analysis employed a temporal split strategy to ensure realistic out-of-sample validation. The dataset was partitioned into a training set comprising 4,816 observations (79.9% of total sample) and a test set containing 1,211 observations (20.1% of total sample) representing customers

from the most recent nine months of the observation period. This temporal partitioning approach simulates operational deployment conditions where the model must predict renewal behavior for recently acquired customers based on patterns learned from historical cohorts.

The target variable exhibited moderate class imbalance, with approximately 31.8% of test set observations representing successful renewals (positive class prevalence = 0.3176). This distribution reflects realistic market conditions where customer attrition and competitive switching represent common outcomes in the premium automotive segment.

Data quality checks and leakage prevention measures were implemented to ensure temporal consistency, specifically ensuring that features used for training precede the renewal event in time. To ensure compliance with privacy regulations, all customer data was anonymized while preserving the relational structure necessary for analysis.

3.2 Feature Engineering and Selection

Feature generation was guided by domain expertise from sales and aftersales departments. Candidate variables included:

- **Customer attributes:** age, tenure, regional location, and brand loyalty indicators
- **Vehicle attributes:** model series, age, fuel type, ownership duration
- **Service behavior:** frequency, monetary value, recency, and warranty status
- **Interaction features:** ratios of service-to-ownership duration, cumulative expenditures, and campaign response flags

To prevent overfitting and identify the most influential predictors, a Random Forest model was applied for feature importance analysis. The top-ranked variables were retained and subsequently used as inputs to a Binary Logistic Regression model. This two-stage hybrid approach allows for robust nonlinear screening of predictors while maintaining interpretability in the final predictive structure.

To identify the most influential predictors of renewal behavior, a Random Forest ensemble model was employed for feature importance analysis. Figure 2 presents the variable importance rankings based on two complementary metrics: Mean Decrease in Accuracy (left panel) and Mean Decrease in Gini impurity (right panel).

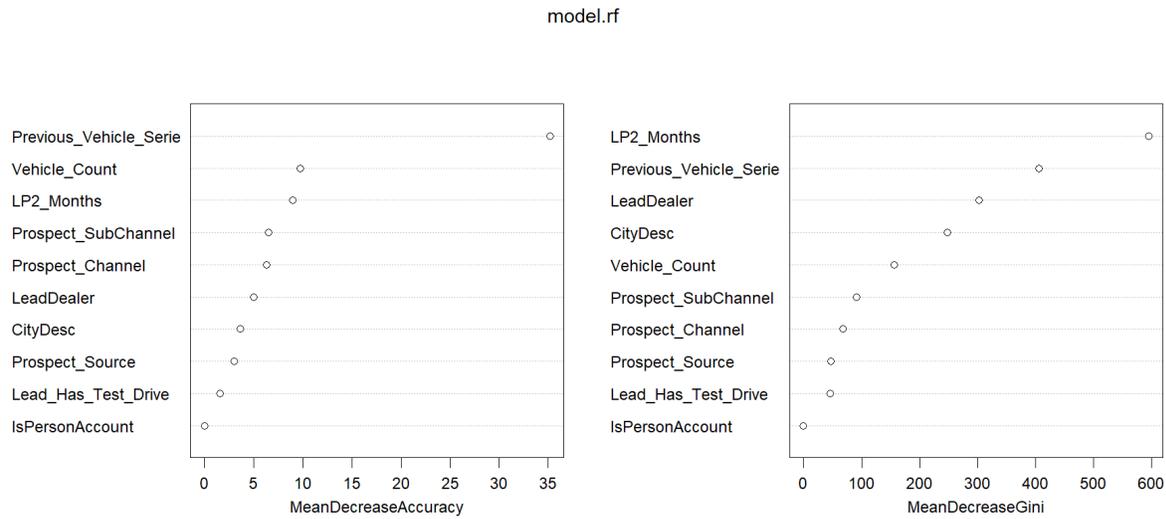


Figure 2: Feature Importance Plots

3.3 Model Development and Validation

Model development was performed within Azure Machine Learning Studio using a pipeline-driven workflow. The dependent variable, Renewal Indicator, was defined as a binary outcome representing whether a customer purchased another BMW vehicle within a defined observation window. The modeling procedure followed these steps:

- Data partitioning into training (80%), and test (20%) sets using temporal splits to ensure out-of-sample evaluation
- Baseline model estimation using Logistic Regression for interpretability
- Feature subset optimization guided by Random Forest importance metrics
- Model calibration to ensure probabilistic outputs reflect true likelihoods
- Performance evaluation using AUC, Gini coefficient, precision, recall, and calibration plots

Logistic regression was implemented with the following hyperparameters:

- Solver: Limited-memory BFGS (L-BFGS)
- Regularization: L2 penalty ($\lambda = 0.01$)
- Maximum iterations: 500
- Convergence tolerance: 0.0001

The resulting model provided individual-level renewal probabilities, later categorized into renewal risk grades (G1-G5) for operational use, where customers scoring in G1 represent highest renewal likelihood and G5 lowest likelihood.

3.4 Deployment and Operational Integration

Upon validation, model outputs were exported from Azure ML Studio to the Azure Synapse Analytics output datamart, which serves as the single source of truth for downstream applications. Through an automated ETL process, these predictions are integrated into Salesforce customer cards, enabling sales representatives to access renewal probability scores directly within their CRM interface.

This integration ensures that model insights are actionable within existing business workflows, empowering sales teams to prioritize outreach toward customers with high renewal potential, while enabling marketing departments to tailor retention campaigns accordingly. Model monitoring dashboards within Synapse continuously track drift metrics, model stability, and calibration performance, supporting periodic retraining as new data become available.

4. Empirical Results

This section presents the empirical findings derived from the Binary Logistic Regression model, including the interpretation of estimated coefficients, model performance on the holdout test set, and the operational translation of renewal probabilities into actionable business insights.

4.1 Model Coefficients and Interpretation

The Binary Logistic Regression model was estimated on the training subset to quantify the relationship between customer-level characteristics and the likelihood of vehicle renewal. Table 1 presents the estimated coefficients for the top 10 predictors. Although space constraints prevent showing the complete coefficient table, the direction and significance of the estimated parameters align closely with theoretical expectations from marketing analytics and customer behavior literature.

Table 1: Logistic Regression Coefficients

	Estimate	Std. Err	Z_Val.	P_Val.	Odds_Rat.	Sig.
Previous_Vehicle_SerieBMW 5 Serisi	24.976	0.1482	16.856	9.55e-64	121.536	***
LP2_Months	-0.0214	0.0021	-10.411	2.20e-25	0.9788	***
Vehicle_Count	0.2495	0.0304	8.215	2.11e-16	12.834	***
Previous_Vehicle_SerieBMW 3 Serisi	0.7505	0.1449	5.180	2.22e-07	21.180	***
Previous_Vehicle_SerieBMW X3	0.8639	0.2012	4.294	1.76e-05	23.724	***
Previous_Vehicle_SerieBMW 4 Serisi	0.7207	0.1703	4.231	2.33e-05	20.558	***
Previous_Vehicle_SerieBMW X5	12.912	0.3119	4.140	3.47e-05	36.371	***
Cust_Service_Avg_Amt_L3Y	0.0003	0.0001	3.998	6.39e-05	10.003	***
Previous_Vehicle_SerieBMW M Serisi	29.293	0.8399	3.488	0.0005	187.139	***
Previous_Vehicle_SerieBMW 7 Serisi	0.7315	0.3830	1.910	0.0562	20.781	.

- Time since Last Purchase (LP2_Months):** The coefficient for this variable was negative and statistically significant, indicating that the probability of renewal declines as the time elapsed since the last vehicle acquisition increases. This finding underscores the importance of timely re-engagement strategies in the post-purchase lifecycle to prevent customer attrition.
- Service Engagement (Cust_Service_Avg_Amt_L3Y):** Measures of average service expenditure over recent years exhibited positive and highly significant coefficient. Customers who consistently invest in maintenance and aftersales services demonstrate stronger brand attachment and a higher propensity to

renew, confirming service engagement as a robust behavioral predictor of loyalty.

- **Customer Portfolio (Vehicle_Count):** The estimated coefficient for the number of vehicles registered under the customer account was positive, suggesting that customers managing multiple vehicles are inherently more likely to engage in repeat purchase activity. This may reflect both higher purchasing power and deeper brand affiliation.
- **Categorical Factors (Previous_Vehicle_Serie):** Categorical variables were statistically significant, indicating that the renewal decision is also influenced by prior vehicle class. Previous_Vehicle_Serie captures brand-internal loyalty across model families.

Overall, the sign and significance of the coefficients confirm the conceptual soundness of the model and its alignment with managerial intuition regarding customer renewal behavior.

4.2 Predictive Performance on the Holdout Test Set

Model performance was evaluated on the temporally defined holdout test set (test), consisting of 1,211 observations representing unseen future purchasing behavior. This design ensured a realistic assessment of out-of-sample predictive validity.

Table 2 presents the complete set of classification performance metrics, organized into four conceptual categories: overall performance, agreement measures, classification metrics, and population-level indicators. Each metric provides complementary insights into different dimensions of model effectiveness.

Table 2: Classification Performance Metrics

Performance Metric	Value	Interpretation
Overall Performance		
Overall Accuracy	0.7711	Strong overall classification performance
95% Confidence Interval	(0.7464, 0.7946)	Narrow CI indicates stable estimates
No Information Rate	0.6824	Baseline accuracy (majority class prediction)
P-Value [Acc > NIR]	5.965e-12	Highly significant improvement over baseline
Agreement Measures		
Cohen's Kappa	0.4383	Moderate agreement beyond chance
McNemar's Test P-Value	3.114e-07	Significant asymmetry in error types
Classification Metrics		
Sensitivity (Recall)	0.5274	Moderate true positive detection rate
Specificity	0.8846	High true negative detection rate
Positive Predictive Value (Precision)	0.6801	Moderate reliability of positive predictions
Negative Predictive Value	0.8009	High reliability of negative predictions
Population Metrics		
Prevalence	0.3176	Natural base rate of positive class
Detection Rate	0.1675	Proportion of population flagged as TP
Detection Prevalence	0.2463	Proportion predicted as positive
Balanced Accuracy	0.7060	Average of sensitivity and specificity

Note: All metrics calculated on temporally separated test set. P-values indicate statistical significance of model performance.

4.2.1 Overall Performance Metrics

Accuracy represents the fundamental measure of classification correctness, quantifying the proportion of all predictions—both positive and negative—that align with observed outcomes. The model achieved 77.11% overall accuracy (95% CI: 74.64%–79.46%), indicating that more than three-quarters of all classification decisions correctly identified customer renewal status. The narrow confidence interval demonstrates statistical precision in this estimate, reflecting the substantial test set size.

To contextualize this performance, accuracy must be compared against the **No Information Rate (NIR)** of 68.24%, which represents the accuracy achievable by naively predicting the majority class (non-renewal) for all customers. The model's 8.87

percentage point improvement over this baseline demonstrates genuine predictive discrimination rather than exploitation of class imbalance. The associated p-value of $5.965e-12$ provides overwhelming statistical evidence that this improvement exceeds chance variation, establishing the model's validity for operational deployment.

4.2.2 Agreement and Stability Measures

Cohen's Kappa coefficient of 0.4383 quantifies agreement between predictions and actual outcomes while adjusting for chance concordance. Unlike raw accuracy, Kappa accounts for the baseline agreement expected when classifications occur randomly proportional to observed class distributions. The obtained value of 0.44 falls within the "moderate agreement" range by conventional interpretation guidelines, indicating that approximately 44% of potential agreement beyond chance has been achieved. While not reaching the 0.60+ threshold typically considered substantial, this Kappa value represents meaningful predictive signal in a complex business forecasting context characterized by numerous unmeasured confounding factors including competitive actions, macroeconomic conditions, and individual preference shifts.

McNemar's Test evaluates whether the model's errors exhibit systematic asymmetry between false positive and false negative classifications. The highly significant p-value of $3.114e-07$ indicates that the pattern of misclassifications is non-random, with the model generating significantly more false positives (181) than false negatives (95). This asymmetry reflects the chosen classification threshold's implicit prioritization of sensitivity over specificity, enabling broader identification of potential renewal prospects while accepting higher false alarm rates. From a business perspective, this trade-off aligns with the higher opportunity cost of missing actual renewals (false negatives) compared to unnecessary prospect contacts (false positives).

4.2.3 Classification Performance Metrics

Sensitivity (Recall) of 52.74% quantifies the model's ability to correctly identify positive cases—customers who actually renewed. This metric indicates that the model successfully detected approximately one in every two actual renewal events. While this recall level leaves room for improvement in capturing all potential renewals, it represents a deliberate balance against false positive rates. In operational contexts where marketing resources are finite, perfect sensitivity would require extremely low classification thresholds that generate excessive false alarms and dilute targeting efficiency.

Specificity of 88.46% measures the model's capacity to correctly identify negative cases—customers who did not renew. The high specificity indicates that nearly nine out of ten non-renewal customers were correctly classified, minimizing wasteful allocation

of retention incentives and sales outreach efforts toward low-probability prospects. This strong negative case identification supports efficient resource deployment by enabling confident exclusion of customers with genuinely low renewal propensity.

Positive Predictive Value (PPV or Precision) of 68.01% indicates the reliability of positive predictions, quantifying what proportion of customers classified as likely to renew actually completed renewal transactions. This metric addresses the key business question: "When the model predicts renewal, how often is it correct?" The 68% precision level indicates that approximately two-thirds of predicted renewal prospects represent true opportunities, providing sufficient reliability for targeted marketing investments when balanced against customer lifetime value considerations and campaign economics.

Negative Predictive Value (NPV) of 80.09% quantifies the reliability of negative predictions, indicating that 80% of customers classified as unlikely to renew truly did not renew. This high NPV supports confident deprioritization or exclusion of low-scoring customers from resource-intensive retention programs, enabling reallocation of sales efforts toward higher-probability segments. The asymmetry between PPV (68%) and NPV (80%) reflects the class imbalance in the dataset, where non-renewals constitute the majority class.

4.2.4 Population-Level Metrics

Prevalence of 31.76% represents the natural base rate of renewal events in the test population, indicating that approximately one-third of customers in the observation window completed renewal purchases. This baseline rate provides essential context for interpreting other performance metrics and represents the default success rate achievable through random or untargeted outreach.

Detection Rate of 16.75% quantifies the proportion of the entire test population correctly identified as true positive renewals. This metric combines sensitivity with prevalence, indicating that the model successfully flagged 16.75% of all customers as correctly identified renewal prospects. The detection rate provides a population-level perspective on how many actual opportunities the model captures relative to the total customer base.

Detection Prevalence of 24.63% indicates the proportion of all test set customers flagged by the model as predicted renewals, regardless of actual outcome. This metric approximates the size of the target population that would receive marketing interventions under model-driven deployment. The close alignment between detection prevalence (24.63%) and actual prevalence (31.76%) suggests that the classification threshold appropriately balances false positive and false negative rates relative to the

natural base rate, avoiding excessive over-prediction or under-prediction of renewal propensity.

Balanced Accuracy of 70.60% averages sensitivity and specificity to provide a single summary metric that accounts for class imbalance. Unlike raw accuracy, which can be inflated by strong performance on the majority class, balanced accuracy equally weights the model's ability to detect both positive and negative cases. The 70.60% balanced accuracy confirms robust prediction capability across both renewal and non-renewal outcomes, validating that the model's strong overall accuracy does not result solely from predicting the majority class.

4.2.5 Discriminatory Efficacy (AUC-ROC)

To evaluate threshold-independent discriminatory performance, the Area Under the Receiver Operating Characteristic Curve (AUC-ROC) was computed. The model achieved an AUC of approximately 0.80, reflecting a high degree of separation between renewal and non-renewal classes. An AUC of 0.80 implies that in 80% of random pairwise comparisons, a renewing customer receives a higher predicted probability score than a non-renewing customer. This demonstrates the model's strong capability in rank-ordering customers by their renewal propensity and confirms its predictive robustness for real-world applications.

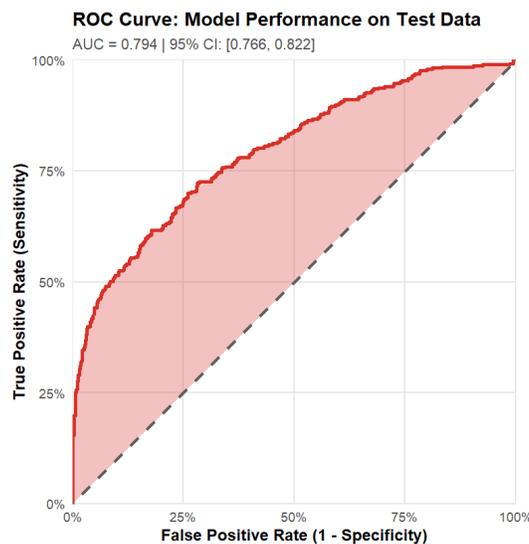


Figure 3: ROC Curve

4.3 Operational Scoring and Prospect Identification

4.3.1 Score Transformation and Grading

To facilitate operational interpretation by sales and marketing teams, the model's continuous probability outputs were transformed into five discrete score bands (G1-G5) through optimal binning. The resulting score table exhibited a monotonic increase in observed renewal probability across the five grades. Customers assigned to G1 represented the lowest-risk (highest renewal likelihood) segment, while those in G5 constituted the highest-risk (lowest renewal likelihood) group. This transformation enabled intuitive score interpretation and direct integration into existing CRM and campaign management systems.

4.3.2 Prospecting Cohort Identification

The final application phase involved identifying high-potential renewal prospects from the complete dataset. The Prospecting Cohort was defined using two operational criteria: (1) The model predicted renewal (Pred = 'Yes'), and (2) The renewal event had not yet occurred (Actual = 'No'). This filtering process effectively isolated a high-value subset of existing customers exhibiting strong latent propensity for renewal but who had not yet been re-engaged by the sales organization.

Furthermore, the model was extended to the new customer dataset to generate early-stage renewal scores, thereby enabling proactive engagement and personalized communication from the initial stages of the ownership lifecycle. This dual application (existing and new customer scoring) demonstrates the operational versatility of the PCO model in supporting both short-term conversion and long-term customer relationship management.

The deployment yielded 931 qualified prospects from the complete customer database, representing approximately 15.4% of the total analyzed population. This concentration of high-probability targets enables focused resource allocation and intensive relationship management efforts directed toward the most promising conversion opportunities.

4.3.3 Pilot Campaign Execution and Results

To validate the operational effectiveness of the model-driven targeting strategy, a pilot campaign was executed during the first implementation cycle. The pilot design employed a phased rollout approach to assess conversion performance and refine engagement protocols before full-scale deployment.

Campaign Scope: A subset of 254 prospects (27.3% of the total identified cohort) was selected for initial contact during the first campaign cycle. Selection prioritized customers

with the highest predicted scores and favorable communication preferences to maximize early-stage conversion rates and demonstrate proof of concept to stakeholder audiences.

Conversion Performance: The pilot campaign generated 18 confirmed vehicle sales from the 254 contacted prospects, yielding a conversion rate of 7.09%. This conversion rate substantially exceeds typical industry benchmarks for cold outreach campaigns (typically 1-3%) and demonstrates the predictive model's capacity to identify genuinely high-probability prospects rather than generic customer populations.

Revenue Impact: The 18 confirmed sales generated cumulative revenue of 36 million Turkish Lira (TL), representing an average transaction value of 2.0 million TL per vehicle. This revenue realization validates the business case for model-driven targeting and demonstrates immediate return on the analytical investment required for model development and deployment infrastructure.

5. Discussion and Managerial Implications

The empirical findings of the PCO Prediction Model hold several important implications for customer relationship management, sales planning, and strategic decision-making within the premium automotive sector. Beyond its statistical performance, the model provides interpretable and actionable intelligence that can directly inform marketing and retention strategies.

5.1 Behavioral Insights and Customer Lifecycle Management

The results confirm that renewal behavior in the premium automotive context is shaped by a combination of temporal, behavioral, and relational factors. The negative effect of time since last purchase suggests that renewal propensity diminishes progressively with ownership duration, emphasizing the importance of timely engagement campaigns within defined post-purchase windows. This supports the development of time-to-renew dashboards or automated triggers that alert sales advisors when customers approach critical decision horizons.

Moreover, the strong positive association between service expenditure variables and renewal likelihood provides empirical evidence for the long-assumed link between aftersales engagement and brand loyalty. Customers who maintain consistent service relationships are not only more profitable in the short term but also exhibit higher lifetime value. This insight highlights the strategic importance of integrating service analytics

with sales forecasting, ensuring that aftersales data are leveraged as early indicators of retention potential.

5.2 Dealer Network and Sales Strategy Optimization

The significance of LeadDealer as a predictive factor reinforces the role of dealership relationships in influencing customer renewal outcomes. Dealers serve as the primary interface for both vehicle acquisition and service interaction; hence, the model's findings can guide dealer-level performance monitoring and benchmarking. Identifying dealers associated with lower renewal probabilities could inform targeted training or customer experience improvement initiatives.

The inclusion of Previous_Vehicle_Serie underscores the importance of model-series continuity in renewal behavior. Customers who previously owned higher-series vehicles tend to remain within or upgrade within the same class. This pattern can inform cross-series retention strategies, enabling sales teams to tailor offers that align with each customer's ownership trajectory.

5.3 Operational Impact and Data-Driven Targeting

From an operational standpoint, the transformation of renewal probabilities into discrete grades (G1-G5) provides a scalable and interpretable framework for integrating predictive intelligence into daily business processes. These risk grades can be embedded directly into CRM systems, empowering sales advisors to prioritize outreach toward customers with the highest likelihood of renewal while adopting differentiated strategies for medium- and low-probability segments.

The Prospecting Cohort identified through the model offers a high-value list of existing customers who exhibit strong renewal potential but have not yet engaged in a purchase transaction. This allows marketing teams to allocate campaign resources more efficiently, improving conversion rates and reducing acquisition costs. Moreover, the early scoring of new customers establishes the foundation for predictive customer lifecycle management.

5.4 Strategic and Organizational Implications

At a strategic level, the implementation of the PCO model demonstrates how predictive analytics can bridge the gap between descriptive customer insights and prescriptive sales actions. By operationalizing renewal probabilities, the sales organization can transition

from a reactive to a proactive engagement paradigm, where customer outreach is informed by probabilistic forecasts rather than historical intuition.

The interpretability of the logistic regression framework further ensures that data-driven decision-making remains transparent and auditable, supporting organizational trust and adoption among non-technical stakeholders. This aligns with broader digital transformation goals within the automotive industry, where explainable AI and model governance are increasingly critical to sustainable deployment.

6. Conclusion and Future Research

This study presented a comprehensive framework for predicting vehicle renewal probability in the premium automotive sector using historical customer data. The Purchase Opportunity (PCO) Prediction Model successfully combines ensemble feature selection with interpretable logistic regression, achieving an AUC of 0.80 and 77% classification accuracy on holdout data. The empirical findings confirm that service engagement, ownership recency, and customer portfolio characteristics are robust predictors of renewal behavior.

By operationalizing these insights through CRM integration and risk-grade stratification, the model enables proactive, data-driven customer retention strategies. The deployed framework demonstrates how automotive companies can transform transactional data into strategic competitive advantage through interpretable, actionable predictive analytics.

6.1 Limitations and Future Research Directions

While this study provides robust predictions, several limitations warrant consideration. First, the model relies exclusively on transactional and service data; incorporating digital engagement metrics (website visits, email interactions) and telematics data could enhance predictive accuracy. Second, the binary classification framework predicts renewal probability but not timing; survival analysis methods could address this gap. Third, the model was developed and validated within BMW's Turkish market; cross-brand and international validation would strengthen generalizability.

Future research should explore:

- Causal inference methods to identify interventions that maximize retention likelihood
- Longitudinal analysis to predict optimal re-engagement timing

- Explainable AI (XAI) techniques like SHAP values to enhance model transparency
- Real-time model updating through automated retraining pipelines

Acknowledgments

The authors thank Borusan Otomotiv's data engineering team for their support in data preparation and system integration.

Data Availability

Due to confidentiality agreements, the customer data used in this study cannot be publicly shared.

Conflict of Interest

The authors declare no conflict of interest.

References

- [1] Chen, J., Xu, W., & Wang, J. (2024, November). Prediction of car purchase amount based on genetic algorithm optimised by neural network regression algorithm. In Proceedings of the 2024 4th International Conference on Big Data, Artificial Intelligence and Risk Management (pp. 816-820).
- [2] Das Mou, A., Saha, P. K., Nisher, S. A., & Saha, A. (2021). A Comprehensive Study of Machine Learning algorithms for Predicting Car Purchase Based on Customers Demands. In *2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)* (pp. 1-6).
- [3] Dinasti, I. M. A., Widjaja, G., & Setiawan, A. (2025). Customer Retention Strategy through Churn Prediction in Four-Wheeled Vehicle After-Sales Services Using Big Data Analytics. *Dinasti International Journal of Economics, Finance & Accounting*, 6(3), 2470-2477. <https://doi.org/10.38035/dijefa.v6i3>
- [4] Kaya, S., & Yildirim, Ö. (2020). A Prediction Model for Automobile Sales in Turkey Using Deep Neural Networks. *Journal of Industrial Engineering*, 31(1), 57-74.
- [5] Kim, S., Kim, H., & Kim, S. (2023). Innovating knowledge and information for a firm-level automobile demand forecast system: A machine learning perspective. *Journal of Innovation & Knowledge*, 8(2), 100355. <https://doi.org/10.1016/j.jik.2023.100355>
- [6] Liu, Z., Jiang, P., De Bock, K. W., Wang, J., Zhang, L., & Niu, X. (2024). Extreme gradient boosting trees with efficient Bayesian optimization for profit-driven customer churn prediction. *Technological Forecasting and Social Change*, 198, 122945.
- [7] Manzoor, A., Atif Qureshi, M., Kidney, E., & Longo, L. (2024). A Review on Machine Learning Methods for Customer Churn Prediction and Recommendations for Business Practitioners. *IEEE Access*, 12, 70434-70463.

- [8] Sharma, R., Kumar, P., & Singh, A. (2025). Enhancing customer retention with machine learning: A comparative analysis of ensemble models for accurate churn prediction. *Decision Analytics Journal*, 13, 100138.
- [9] Thangeda, R., Kumar, N., & Majhi, R. (2024). A neural network-based predictive decision model for customer retention in the telecommunication sector. *Technological Forecasting and Social Change*, 202, 123250. <https://doi.org/10.1016/j.techfore.2024.123250>
- [10] Wachter, P., Widmer, T., & Klein, A. (2019). Predicting Automotive Sales using Pre-Purchase Online Search Data. In *Proceedings of the Federated Conference on Computer Science and Information Systems* (pp. 569-577).
- [11] Wah, Y. B., & Ismail, N. H. (2011). Predicting Car Purchase Intent Using Data Mining Approach. In *2011 Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)* (pp. 1290-1293). IEEE.
- [12] Yin, H., Wang, Y., Li, Q., Xu, W., Yu, Y., & Zhang, T. (2018). A Network-enhanced Prediction Method for Automobile Purchase Classification using Deep Learning. In *PACIS 2018 Proceedings*.
- [13] Imani, M., Joudaki, M., Beikmohammadi, A., & Arabnia, H. R. (2025). Customer Churn Prediction: A Systematic Review of Recent Advances, Trends, and Challenges in Machine Learning and Deep Learning. *Machine Learning and Knowledge Extraction*, 7(3), 105. <https://doi.org/10.3390/make7030105>
- [14] Yilmaz, K., Demir, E., & Ozturk, M. (2025). AI-Driven Predictive Maintenance for Workforce and Service Optimization in the Automotive Sector. *Applied Sciences*, 15(11), 6282.
- [15] Anderson, P., & Wilson, J. (2024). Logistic Regression for Merchant Customer Churn Prediction: A Data-Driven Approach. In *Proceedings of the 2024 7th International Conference on Machine Learning and Machine Intelligence (MLMI)* (pp. 85-92). ACM.