International Journal of Emerging Trends in Computer Science and Information Technology
ISSN: 3050-9246 | https://doi.org/10.63282/3050-9246/ ICRTCSIT-110
Eureka Vision Publication | ICRTCSIT'25-Conference Proceeding

Original Article

Optimizing Direct Marketing For Burial Insurance With Machine Learning

Mr. Vaibhav Tummalapalli
Data science manager, Epsilon data management LLC.

Abstract - Burial insurance, a niche but vital segment of life insurance, presents persistent marketing challenges due to low consumer awareness, price sensitivity, and heterogeneous response behavior. Despite widespread use of propensity modeling in broader insurance marketing, there is limited published research applying advanced machine learning approaches to burial insurance acquisition campaigns. This study addresses this gap by developing and validating a predictive modeling framework to improve campaign targeting for burial insurance products. Using a representative sample drawn from a production database of approximately 1.3 million consumer records with historical burial insurance response labels and enriched demographic, lifestyle, psychographic, and financial attributes, we identified key predictors of responseincluding low income, high mobility, limited assets, and behavioral markers such as sweepstakes participation. Ensemble models such as Gradient Boosting and XGBoost delivered a strong performance, achieving 1.8x lift in the top decile and capturing 43% of total responders in the top three deciles. These results demonstrate the value of data-driven segmentation and targeting in reducing acquisition costs and improving ROI for burial insurance marketing, with implications for applying machine learning approaches to other underserved, price-sensitive insurance markets.

Keywords - Predictive/Propensity modeling, Machine Learning in Marketing, Insurance Marketing, Customer Segmentation, Burial Insurance.

1. Introduction

Burial insurance, also known as final expense insurance, is a niche life insurance product designed to cover end-of-life costs such as funeral expenses, outstanding medical bills, and other final obligations. It is primarily marketed to older adults and individuals with limited financial resources who may not qualify for or afford traditional life insurance products. Typical target segments include seniors, renters, lower-income households, and those on fixed incomes populations marked by financial vulnerability, budget constraints, and heightened concern about leaving final expenses for family members. These characteristics demand highly tailored marketing strategies that not only identify likely buyers but also understand their underlying motivations, needs, and concerns[1]. Unlike broader life insurance products sold through agents or digital channels, burial insurance relies heavily on direct mail to reach older, lower-income, and often geographically mobile segments with smaller premiums that reduce agent incentive necessitating distinct segmentation and outreach strategies tailored to these unique operational and ethical challenges

Marketing burial insurance poses distinctive acquisition challenges beyond high customer acquisition costs. Consumers must trust an insurer's intangible promise of future payment and navigate product complexity without extensive personal financial expertise[5,6]. In traditional channels, intermediaries like agents bridge this gap by advising clients and personalizing coverage. However, direct marketing channels such as direct mail lack this interpersonal support, making it more difficult to explain complex benefits and build trust with financially cautious or risk-averse consumers. This complexity makes it essential not only to target the right consumers efficiently but also to generate meaningful insights into their preferences, constraints, and potential objections insights that can inform both marketing strategy and messaging design.

Direct mail remains a dominant acquisition channel for burial insurance because it complies with strict marketing regulations and effectively reaches older, geographically dispersed populations[9]. Yet it also faces persistent challenges: low baseline response rates, high printing and distribution costs, and difficulty in personalizing offers at scale. These limitations underscore the importance of precise segmentation and targeting strategies to maximize return on investment while also enabling creative teams to tailor messaging that resonates with specific consumer profiles. Advanced segmentation can help marketers move beyond broad demographic lists to develop offers that address identified needs, preferences, and sensitivities.

Foundational marketing theory emphasizes segmentation, targeting, and positioning (STP) as essential pillars of effective strategy. Segmentation divides heterogeneous markets into more homogeneous groups with shared characteristics, enabling marketers to tailor value propositions, communications, and even product features more effectively[2,3]. In insurance marketing, segmentation is critical not only for aligning product design with customer needs but also for meeting regulatory expectations and ensuring ethical outreach to vulnerable consumer groups . Ethical outreach involves offering appropriate, transparent, and fair marketing that avoids exploiting informational asymmetries or financial distress, while also supporting product development that aligns with consumer capabilities and values.

Predictive modeling has emerged as a crucial tool in direct marketing to improve targeting precision and ROI by identifying high-propensity segments[7] analyzed the challenges of customer acquisition modeling and demonstrated how response scoring approaches can prioritize prospects based on predicted likelihood to enhance marketing profitability. Database marketing frameworks emphasize the role of statistical and machine learning models in optimizing customer acquisition, improving retention, and increasing cross-selling effectiveness[4]. While much research on predictive modeling in insurance has focused on retention and churn reduction[11], acquisition modeling particularly for niche products like burial insurance has received less academic attention. There is also limited published work on how modeling insights can support creative strategy, inform product refinement, and enable more coordinated omnichannel planning in the insurance context. This gap reflects several barriers: proprietary customer lists rarely shared for research, regulatory and privacy constraints limiting data access, and an industry focus on broader life insurance markets over specialized, lower-premium burial products that primarily target financially vulnerable consumers

Moreover, ethical considerations are paramount when marketing to older, lower-income, and financially vulnerable consumers. Segmentation strategies should not only increase targeting efficiency but also enhance consumer choice, reduce wasteful and irrelevant outreach, and mitigate the risk of predatory or inappropriate marketing practices[8,10]. Incorporating predictive modeling into these strategies supports more transparent, evidence-based decision-making that can be updated and refined as consumer behavior and economic conditions evolve.

This study addresses that gap by developing and validating a machine learning–based propensity modeling framework specifically for burial insurance direct marketing campaigns. Using a representative sample drawn from a production database of approximately 1.3 million consumer records enriched with demographic, lifestyle, psychographic, and financial variables, the analysis seeks to uncover behavioral patterns and response predictors unique to this market. By applying and evaluating ensemble learning techniques such as Gradient Boosting and XGBoost, this research aims not only to enhance targeting precision and reduce acquisition costs but also to generate actionable customer insights, support creative messaging and product design strategies, inform omnichannel planning, and enable future adaptability through ongoing model refinement. The modeling framework produces individualized propensity scores by integrating multiple demographic, geographic, and behavioral variables, enabling marketers to prioritize outreach to consumers with the highest predicted likelihood of response and supporting nuanced, multidimensional segmentation strategies. These improvements can help insurers balance business goals with consumer protection by delivering more relevant, respectful, and effective marketing while minimizing unwanted or inappropriate outreach. The findings also contribute to the broader literature on insurance marketing analytics, segmentation theory, and the ethical practice of marketing to price-sensitive, underserved insurance markets.

2. Data Overview

The dataset used for this study was derived from a comprehensive production database containing approximately 1.3 million consumer records with historical burial insurance campaign response labels, aggregated annually from over 30 distinct public and commercial data sources. These sources include voter registration files, U.S. Census statistics, property deeds and county assessor records, product registration databases, questionnaire responses, mail order purchase behavior, and deceased file listings. The integration of these diverse sources enables the construction of rich, multidimensional consumer profiles essential for accurate targeting in burial insurance marketing. For this analysis, a random sample of approximately 158,000 records was selected to ensure computational feasibility while preserving representativeness of the broader target population. All records were attribute-level only, with personally identifying information removed by the data vendor to maintain privacy and ensure compliance with applicable ethical sourcing standards.

Each consumer record includes a wide array of variables grouped into demographic, lifestyle, psychographic, financial, behavioral, and geographic dimensions. These variables are selected for their relevance to predicting insurance purchasing behavior, especially among older, lower-income, and price-sensitive consumersa key target segment for burial insurance products.

Key data components include:

- Demographics: Age, gender, homeownership, length of residence, household composition, ethnicity
- Financials: Estimated income, mortgage status, net worth, credit behavior
- Lifestyle & Psychographics: Hobbies (e.g., sweepstakes entry, coin collection), value-conscious shopping behavior, mobility scores
- Health and Triggers: Flags for age-related services, Medicaid qualification, Medicare Part D usage
- Geographic & Market Indicators: Designated Market Areas (DMAs), county size, historical response rates by region
- The target variable, labeled RESPONDER, is binary:
- 1 = responders (18,891 records)

- 0 = non-responders (140,461 records)

This yields an overall observed response rate of approximately 11.8%, reflecting the typical challenge of low baseline responsiveness in direct marketing for insurance products. Such class imbalance underscores the need for accurate predictive modeling to identify high-propensity individuals effectively, reduce marketing waste, and improve overall campaign ROI.

This dataset, with its comprehensive feature set and realistic class distribution, provides a robust foundation for developing machine learning models capable of differentiating likely responders from unlikely responders with greater precision

3. Methodology

3.1. Data Cleaning

The initial phase of the modeling pipeline focused on data cleaning to ensure quality, consistency, and analytical readiness. This process included deduplication, structural validation, missing data handling, and categorical simplification. This step was essential for producing marketing insights that can effectively inform burial insurance targeting strategies.

- **Duplicate Removal and Type Validation:** The dataset was first examined for duplicate records to eliminate redundant observations that could bias model learning. All duplicate rows were removed to maintain the integrity and uniqueness of each consumer record. Columns were then reviewed for data type consistency, with each field assigned an appropriate type (numeric, categorical, or string) based on its underlying semantics and intended role in analysis. This validation ensured the data structure was compatible with downstream modeling requirements.
- **Low-Variability and High-Missing Fields:** Features with extremely low variability those containing only a single unique value across all records were excluded, as they offer no predictive value in distinguishing responders from non-responders. For missing data, a threshold-based filtering strategy was implemented. Variables with more than 30% missing values were removed from the dataset. This 30% threshold was chosen to balance retaining valuable predictors with avoiding excessive noise or bias from imputation. Notably, the majority of fields in this dataset exhibited less than 5% missingness, making this a conservative and appropriate choice for ensuring data completeness.
- **Category Consolidation**: Certain categorical variables displayed high cardinality with overly granular subgroupings that could hinder model interpretability and practical marketing use. For example, this variable variablea proprietary segmentation that reflects consumer lifestyle, interests, and socio-economic traits contained highly specific categories (e.g., A1, A2, A3, ..., Z). These were consolidated at a broader grouping level (e.g., all "A" series grouped as "A") to reduce dimensionality while preserving meaningful differentiation. This approach enhances the model’s ability to generate interpretable and actionable marketing segments. A similar consolidation was applied to a financial indicator, a proprietary metric ranking households by purchasing power. Granular score values were converted into broader bands to retain predictive signal while simplifying interpretation for marketing strategy development.
- **Formatting and Standardization**: To ensure data uniformity and analytical readiness, selected fields were reformatted. For instance, the *Home Market Value* variable, initially stored as alphanumeric text (e.g., "$250K+"), was converted into standardized numeric currency values. This transformation enabled consistent comparability across records and supported more precise modeling of financial indicators relevant to insurance purchase propensity. These data cleaning procedures established a high-quality, structured foundation for subsequent exploratory analysis, feature engineering, and predictive modeling. By ensuring the dataset was both analytically robust and business-relevant, this phase laid the groundwork for developing models capable of accurately identifying high-propensity burial insurance prospects, thereby supporting more efficient and effective marketing campaigns.

3.2. Exploratory Analysis

The exploratory data analysis (EDA) phase provided valuable insights into response behaviors across key demographic, geographic, and psychographic variables. By understanding these relationships, we ensure that the model’s segmentation of high- and low-propensity responders is grounded in interpretable, business-relevant drivers. These insights were not only useful for model interpretation but also informed critical preprocessing and feature engineering decisions:

- **Income & Value score trends:** Exploratory analysis revealed a clear inverse relationship between household income levels and observed response rates. As household income bands increased, average response rates declined steadily, ranging from approximately 14% in the lowest band to around 8% in the highest band (*see Figure 1 Response Rate by Income*). A similar pattern emerged in the Value Score variable, which acts as a proprietary proxy for household financial strength. Records in the “Below Average” and “Poor” categories exhibited response rates of approximately 13%, compared to only 9% for the “Best Profit” segment (*see Table 1-Response Rate by Value Score*). These descriptive patterns suggest that lower-income and lower-value-score households tend to respond at higher rates to burial insurance marketing campaigns. While these univariate rates do not control for other factors, they provide preliminary evidence of segmentation potential and informed the inclusion of income and value score features in the predictive modeling pipeline

Table 1. Response Rate by Value Score

Value Score	Response Rate
A : Best profit	9%
B : Above Average	11%
C : Average	12%
D : Below Average	13%
E : Poor	13%

In Fig 1. Bands 1–9 represent rank-ordered income groups from lowest (1) to highest (9); exact thresholds are proprietary

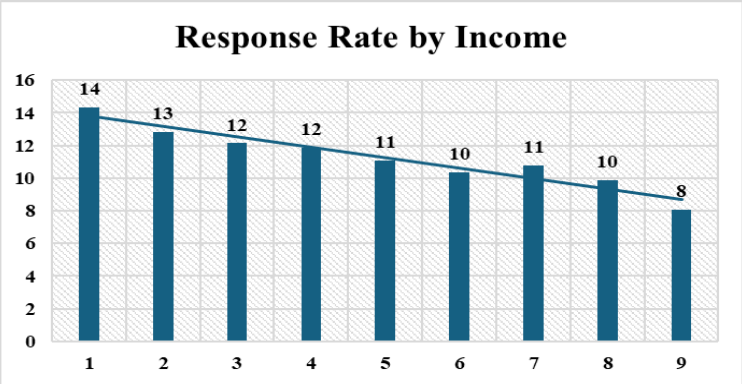


Figure 1. Response Rate by Income

- **County Size & Geography:** Analysis revealed a modest relationship between county size and response rates. Respondents from smaller counties (defined as those outside the largest metropolitan statistical areas) exhibited slightly higher response rates (approximately 13%) compared to respondents in the largest MSAs (approximately 11%) (*see Fig 3. Response Rate by County Size*). These observed differences suggest geographic segmentation effects that may be relevant for targeting.

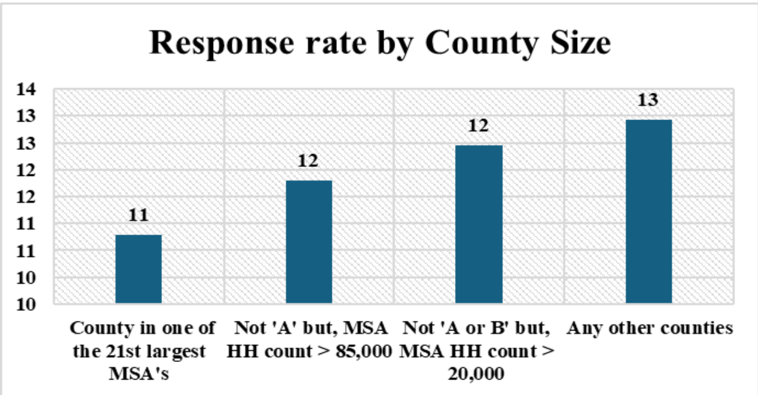


Figure 2. Response Rate by County Size

- **Homeownership & Mobility**: Analysis showed that renters responded at higher rates (15–14%) than homeowners (13%), and individuals with shorter residence durations also exhibited elevated response rates (*see Table 2 and Fig 2.*). These patterns informed the inclusion of homeownership status and mobility measures as segmentation features in the predictive model

Table 2. Response Rate by Homeownership

Home Owner	Response Rate
Unknown	14%
Definite Renter	14%
Probable Renter	15%
Definite Owner	13%
Probable Owner	11%

Length of Residence was coded into 8 ordinal bands, with Band 1 indicating shortest tenure (recent movers) and Band 8 representing longest tenure (most stable households). Exact cutoff thresholds reflect ascending length of stay

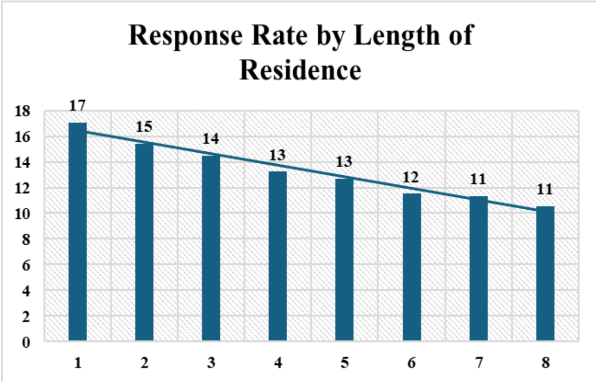


Figure 3. Response Rate by Length of Residence

- **Primary vs Secondary Household Members:** Primary members of the household had a 12% response rate, compared to only 9% for secondary members (See Table-3). This difference suggests that household role is a meaningful segmentation variable and supports including primary/secondary member status as a predictive feature in the modeling pipeline.

Table 3. Response Rate by Member code

Member Code	Response Rate
Primary	12%
Secondary	9%

- **Language & Cultural Considerations:** The data also revealed language-based segmentation patterns:
  - English-speaking households made up 91% of the sample and had a 12% response rate.
  - Spanish-speaking (Hispanic) households represented just 3% of the population and had a slightly lower response rate of 9%.

Although the Hispanic segment is underrepresented in this dataset, these results suggest language preference is a relevant segmentation dimension and support including language flag variables in predictive modeling to capture potential differences in marketing responsiveness.

- **Niche Segments & Outliers:** The dataset includes an internal *Niche Segment* variable that classifies households into 26 lifestyle and behavioral profiles (labeled A–Z). Each segment reflects a unique combination of socio-economic, psychographic, and consumption traits used for marketing segmentation. Certain niche segments (e.g., Groups A and F) as shown in Fig 4 exhibited unusually high response rates (~20%). While most segments clustered around 10–13%, a few smaller segments exhibited higher rates exceeding 20%, which may reflect sample-size effects or specialized profiles. These outliers were reviewed for potential consolidation in modeling to ensure stability.

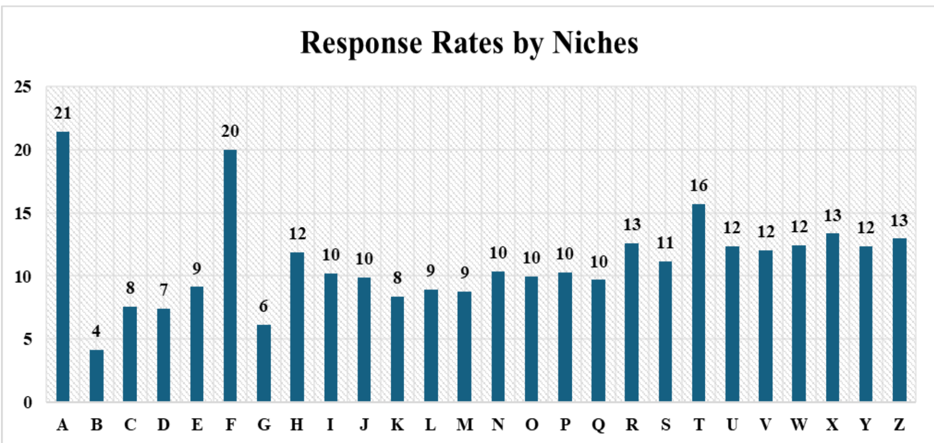


Figure 4. Response Rate by Niches

- **DMA Response Rates & Modeling Constraints:** Geographic analysis at the DMA (Designated Market Area) level revealed high variability in response rates. Notably, a certain DMA showed a response rate of 67%, (See Fig 5.) but this was driven by only three records highlighting the importance of outlier management. While DMA is a. Significant variable, directly encoding each unique DMA would create a high-dimensional feature that could hinder model performance. To address this, target encoding was applied later in the pipeline, where each DMA was replaced with its average response rate derived from the training set allowing us to preserve geographic insight without inflating the feature space

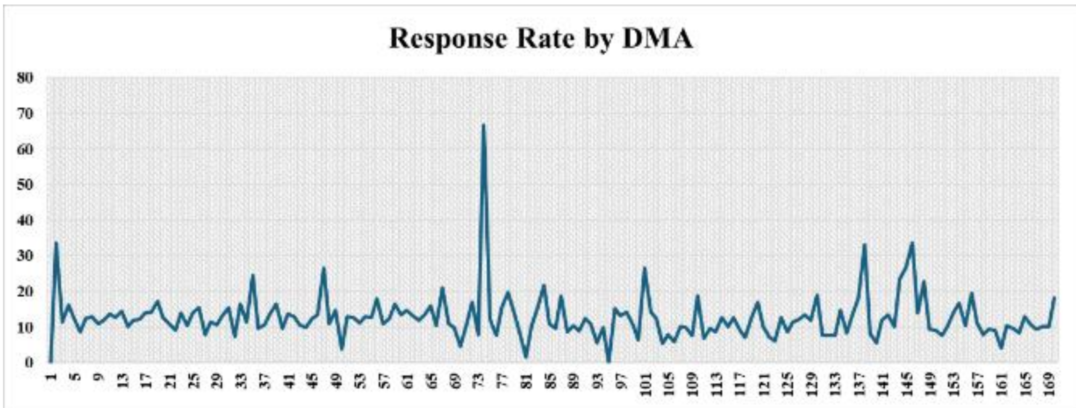


Figure 5. Response Rate by DMA (Designated Market Area)

3.3. Data Preprocessing

Effective preprocessing is a critical step in the modeling pipeline, as it ensures the integrity and quality of the data before feeding it into machine learning algorithms. The following steps were taken after initial data cleaning and exploration analysis:

- **Train-Test Slit:** The dataset was partitioned into training and testing subsets to enable unbiased evaluation of model performance. The train-test split ratio was set **at** 70-30, based on dataset size and computational considerations. While larger datasets (millions of records) may permit a 90-10 split to maximize learning, for moderately sized datasets such as this (~159,352), a 70-30 split strikes a good balance between model training and performance evaluation
- **Missing Value Imputation:** Missing values were addressed using appropriate imputation techniques tailored to the variable types and missingness pattern:
- **Numeric Variables:** Since most numeric fields had a low percentage of missing data (1%–5%), the median was used for imputation. Median is preferred over mean as it is robust to outliers and preserves the distribution shape in skewed data.
- **Categorical Variables:** For categorical features with minimal missingness, missing values were replaced with a constant placeholder value (e.g., "U" for Unknown) or treated as a separate category. This allows the model to learn from the presence or absence of information.

Given the overall high completeness of the dataset, advanced imputation methods such as K-Nearest Neighbors or regression-based techniques were considered but ultimately deemed unnecessary. Simpler methods offered similar efficacy while preserving interpretability and reducing computational overhead.

3.3.1. Categorical Variable Encoding:

Proper encoding of categorical features was essential to prepare the data for modeling:

- One-Hot Encoding (Dummies): For low-cardinality variables, standard one-hot encoding was applied. This ensures each category is represented as a binary feature, preserving interpretability and compatibility with tree-based models.
- Grouped Encoding: High-cardinality categorical variables such as niche segments and value scores were grouped based on shared behavioral traits or business logic (e.g., consolidating all "A" categories into one group). After grouping, dummy indicators were created for the consolidated classes.
- Target Encoding for DMA: The DMA (Designated Market Area) variable had a high number of unique categories and could not be directly one-hot encoded due to the risk of dimensional explosion. To preserve geographic insight without inflating feature space, target encoding was employed:
- For each DMA, the average response rate from the training data was computed.
- These averages were used to replace the DMA codes in both the test and future production datasets, ensuring no data leakage or information bleeding across partitions.
- This technique captures the predictive power of geography while maintaining model scalability

- To avoid data leakage, all target encoding averages were computed strictly within the training data partition. These encodings were then applied to both the test set and any future production data, ensuring no information from the holdout set influenced training

Overall, these preprocessing steps balanced predictive accuracy with interpretability, ensuring the final modeling pipeline would be both technically robust and actionable for marketing practitioners.

3.4. Feature Selection

Feature selection plays a critical role in improving model performance, stability, and interpretability by identifying the most predictive attributes while reducing noise and computational complexity. It is important that marketing teams can understand and trust the drivers of predicted response. A hybrid approach combining statistical and machine learning (ML)-based methods was adopted to ensure robust selection of relevant features[12].

- **Statistical Feature Selection:** Statistical techniques were first applied to perform an initial screening, particularly useful when working with a large number of input features. This approach helps quickly eliminate non-informative features before more computationally intensive modeling steps.
- **Categorical Features**: The Chi-square test was used to evaluate the association between each categorical variable and the target response. Features with the highest chi-square scores were retained.
- **Numerical Features**: For continuous variables, Analysis of Variance (ANOVA**)** was used to determine whether the mean of the response varied significantly across different values of each predictor.
- **Selection Criteria**: The top 20–30 features from each test were shortlisted based on ranking metrics. While other metrics such as Mutual Information (MI) or Information Value (IV) are also viable, exploratory analysis suggested they yielded largely overlapping top-ranked variables in this context.

These shortlisted features were then passed into machine learning–based feature selection models for further refinement.

- **Machine Learning-Based Selection:** Machine learning methods offer a more advanced layer of feature selection by accounting for nonlinear relationships and interactions between variables.
- **Random Forests (RF)** and **Gradient Boosting (GBM)**: These tree-based ensemble algorithms inherently provide feature importance scores. Variables contributing most to reducing error across trees were ranked higher and retained.
- **Recursive Feature Elimination (RFE)**: RFE is a wrapper-based method that iteratively eliminates the least important features based on a model's performance (often using RF or GB as the base estimator). Although effective, RFE can be computationally intensive, especially on large datasets, and was selectively applied for experimentation and cross-check the ranking of top features rather than as a primary selection tool.

This layered approach ensured that the final feature set retained only the most predictive, business-relevant variables while maintaining interpretability and computational feasibility:

- **Dimensionality Trade-Offs and Iteration:** The feature selection process was iterative and heuristic-driven process. Multiple combinations of top-ranked features from both statistical and ML methods were evaluated through cross-validation. The goal was to strike a balance between model simplicity and interpretability, and maximizing predictive performance on unseen data.

By trimming redundant or noisy features, the final model became more efficient, less prone to overfitting, and easier to explain an important consideration for stakeholders in regulated industries like healthcare and insurance

3.5. Model building & Tuning

To develop a robust predictive model for burial insurance campaign response, two high-performing machine learning algorithms were selected: Gradient Boosting Machine (GBM) and Extreme Gradient Boosting (XGBoost**)**. Both are ensemble-based techniques that have consistently demonstrated superior performance on structured, tabular data and are widely adopted in both industry and competitive data science (e.g., Kaggle).

- **Choice of Algorithms and Justification**: Gradient Boosting and XGBoost belong to the family of boosting algorithms that sequentially build decision trees where each new tree attempts to correct the errors of the previous one[13]. This iterative error-reduction process leads to a strong composite model capable of capturing complex, non-linear relationships and feature interactions. From a conceptual standpoint, this is akin to assembling a team of experts each tree contributes incremental knowledge, correcting the missteps of its predecessor.

Their ability to handle mixed data types, manage missing values internally, and incorporate built-in feature importance makes these models particularly well-suited for marketing applications like campaign response modeling.

3.5.1. Training Strategy and Hyperparameter Tuning

Both models were trained on the preprocessed training dataset using a grid search approach with cross-validation to tune hyperparameters. This cross-validation ensured robust parameter selection by balancing model complexity and predictive performance within the training data. The model's performance was then evaluated on a separate holdout test set to confirm generalizability and prevent overfitting.

The hyperparameters tuned for **GBM** included:

- **Learning Rate (Shrinkage Factor):** Controls how much each tree influences the final prediction.
- **Number of Trees:** Total number of boosting iterations.
- **Loss Function:** Log Loss, used for binary classification.
- **Tree Depth:** Controls the complexity of individual trees.
- **Min Samples Split:** Minimum number of samples required to split a node.
- **Max Features:** Number of features considered when looking for the best split.

For XGBoost, in addition to the above, the following regularization parameters were also fine-tuned:

- **Lambda (L2 Regularization):** Penalizes overly complex models to improve generalization.
- **Gamma:** Minimum loss reduction required to make a further split on a leaf node; helps in pruning.

These hyperparameters were optimized to balance model complexity, reduce variance, and maximize predictive accuracy. Final models were selected based on a combination of lift in top deciles, capture rate, and consistency across training and test datasets.

3.6. Results & Model Metrics

To evaluate the effectiveness of the burial insurance response model, we used three core metrics commonly adopted in marketing analytics: Lift, Response Rate by Decile, and Cumulative Capture Rate. These metrics help assess both the predictive power and business value of the model.

Lift measures how much better the model is at identifying responders compared to a random sample. A lift of 1.78 in the top decile indicates that individuals in this group are **1.78 times more likely to respond** than randomly selected individuals. Ideally, the lift should decline monotonically across the deciles, which signals good model calibration. As shown in the first panel of Figure 1, the lift curve for both training and test sets follows a smooth decreasing trend without spikes, suggesting model robustness and lack of overfitting.

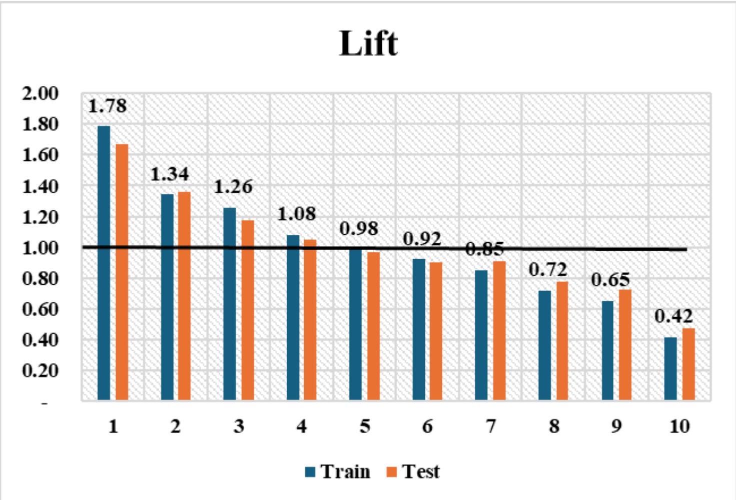


Figure 6. Lift Chart – Train & Test

- **Response Rate by Decile**: gives a direct view of how responses are distributed across the ranked scores. In this case, the top decile has a 21% response rate in training and 20% in test, significantly above the overall baseline (11.8% - which is the overall response rate in the dataset). Decile-wise response rates also show a monotonic decrease, reinforcing model consistency across datasets.

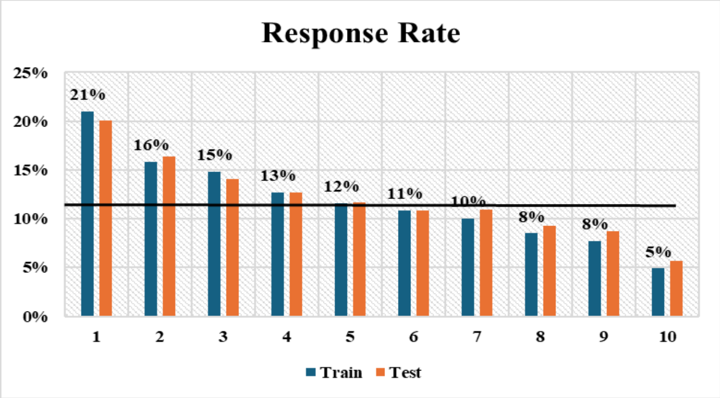


Figure 7. Response Rate by Decile – Train & Test

- **Cumulative Capture Rate**: represents the cumulative percentage of total responders[12] captured as we move through the deciles. The model captures 43% of all responders within the top three deciles (See Fig 8.), demonstrating its ability to prioritize high-value targets effectively. The test curve closely follows the training curve and significantly outperforms the random baseline (green line), indicating high generalizability.

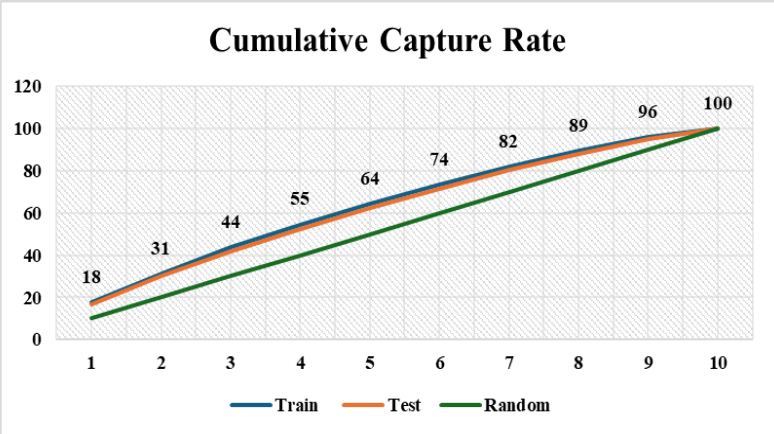


Figure 8. Cumulative Capture Rate – Train & Test

These results confirm the model's stability and predictive strength:

- The top decile alone accounts for 18% of all responders, a 1.78x improvement over random targeting.
- The first three deciles capture nearly half of all responses, enabling efficient campaign targeting.
- The performance is consistent across training and test datasets, confirming the model's ability to generalize.

While these results are strong, it is important to note that the model was developed using a subset of available attributes. With access to the full feature set over 1,400 variables further improvements in predictive accuracy and segmentation depth are anticipated.

3.7. Variables in the Model & Insights

Table 4. Feature Importance Table

Rank	Feature	Feature Importance Scores
1	DMA	0.50
2	Amazon Prime Consumers	0.12
3	Age	0.11
4	Mobility Score	0.08
5	Sweepstakes	0.06
6	Medicaid Qualified Household (PM)	0.04
7	Value Chain Enthusiast (PM)	0.034
8	Coins Collector (PM)	0.030
9	Medicare Plan D Prescription Purchaser (PM)	0.01
10	Income Index	0.003
11	Mortgage (PM)	0.002

To improve interpretability of the model's decision-making process, we analyzed both the relative importance and marginal effects of the top predictive variables. Feature importance was extracted from the Gradient Boosting model using gain-based scoring to identify which features contributed most to prediction accuracy.

To further understand these relationships, partial dependence plots (PDPs)[14] was generated for the most influential variables. These plots display the model's predicted probability of response as a function of a single feature (x-axis), showing the average predicted propensity score (y-axis), while averaging over the distribution of all other features in the data. By visualizing the marginal effect of each input variable on its continuous scale, PDPs reveal how predicted response likelihood changes smoothly across the range of observed values. This approach offers clear insights into variable influence, directionality, and potential non-linear effects, supporting model transparency and actionable interpretation

Historical DMA Response Rate emerged as the most influential predictor. Individuals residing in Designated Market Areas (DMAs) with historically higher campaign response rates were significantly more likely to respond. This regional behavior may reflect localized market saturation, media exposure, and socioeconomic conditions.

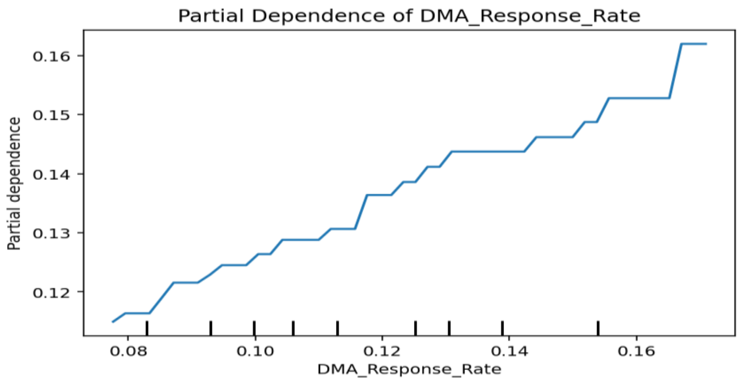


Figure 9. Partial Dependence Plot – DMA

Amazon Prime Consumers is a proprietary vendor-derived ranking from 0 (highest likelihood of being an Amazon Prime subscriber) to 99 (lowest). This score is constructed using demographic and behavioral predictors to estimate a consumer's propensity for premium membership.

Analysis revealed that the Amazon Prime Consumer Score is *negatively associated with burial insurance response rates* individuals with lower scores (i.e., those more likely to be Prime subscribers) were less likely to respond.

Notably, the Amazon Prime Consumer Score is also significantly negatively correlated with the Income Index (Spearman $\rho = -0.475$, $p < 0.0001$), suggesting that households with higher Prime membership likelihood tend to be higher-income and brand-loyal.

This relationship aligns with our earlier EDA findings showing that lower-income households exhibited higher response rates to burial insurance marketing. Together, these results support the interpretation that burial insurance products appeal more strongly to price-sensitive, lower-to-moderate-income segments who are less likely to hold premium retail subscription.

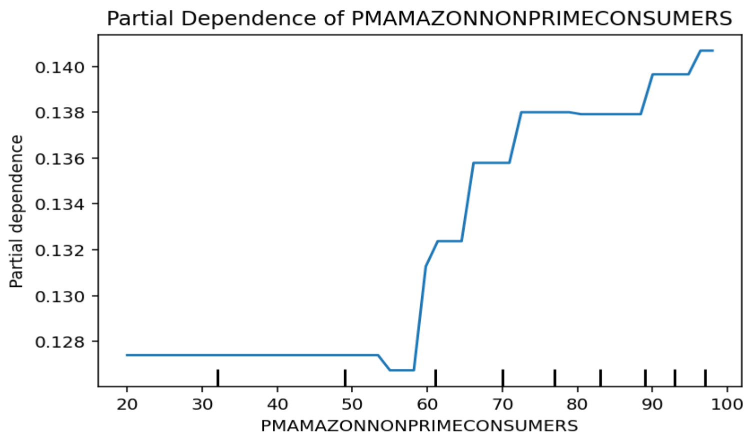


Figure 10. Partial Dependence Plot – Amazon Prime Consumers

Age was positively correlated with response likelihood, consistent with the nature of burial insurance being a product typically marketed to individuals aged 65 and above. We see an increase in response propensity, especially around 60.

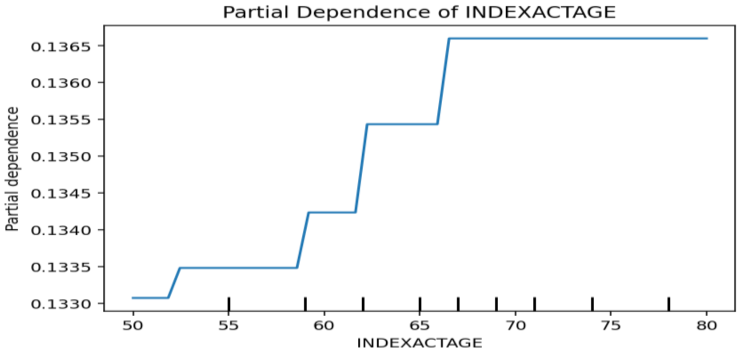
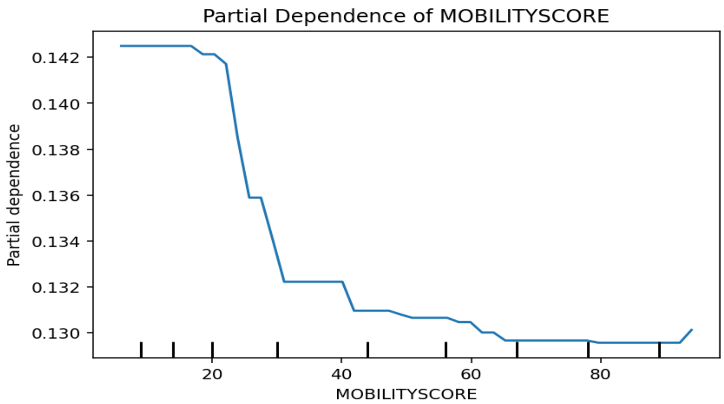


Figure 11. Partial Dependence Plot – Age

Mobility Score is a proprietary vendor-derived ranking from 0 (most likely to move) to 99 (least likely to move), designed to estimate household residential mobility propensity. In the model, lower Mobility Scores were associated with higher predicted response probabilities. This relationship is consistent with exploratory analysis showing higher response rates among renters and households with shorter lengths of residence groups typically characterized by greater residential mobility. These findings suggest that burial insurance marketing may be particularly effective when targeted toward more mobile, potentially financially constrained households seeking accessible coverage options



.

Figure 12. Partial Dependence Plot – Mobility Score

Sweepstakes Participation is a behavioral indicator derived from consumer data capturing whether an individual has a history of entering promotional sweepstakes or similar contests. In the model, this variable was positively associated with higher predicted response probabilities.

This pattern suggests that sweepstakes participation may serve as a proxy for price-sensitive, deal-seeking behaviora relevant trait for targeting burial insurance products, which often appeal to cost-conscious consumers seeking affordable coverage options. Including this variable helps identify segments more receptive to promotional offers and tailored messaging that emphasizes value and accessibility.

Table 5. Response Rate by Sweepstakes

Sweepstakes	Response Rate
N	11%
Y	14%

*Y = Yes, Participates in Sweepstakes; N = No, does not participate in Sweepstakes

Medicaid Qualified Household (PM**)** is a proprietary vendor-derived ranking from 0 (highest likelihood of qualifying for Medicaid) to 99 (least likely). Analysis revealed that lower Medicaid scores were associated with higher predicted response probabilities in the model. This finding is consistent with the product's target audience of older, income-constrained households. Additionally, the Medicaid score showed a significant positive correlation with Income Index (Spearman

$\rho = 0.465$, $p < 0.001$), supporting its use as a proxy for financial need in identifying segments more receptive to burial insurance marketing.

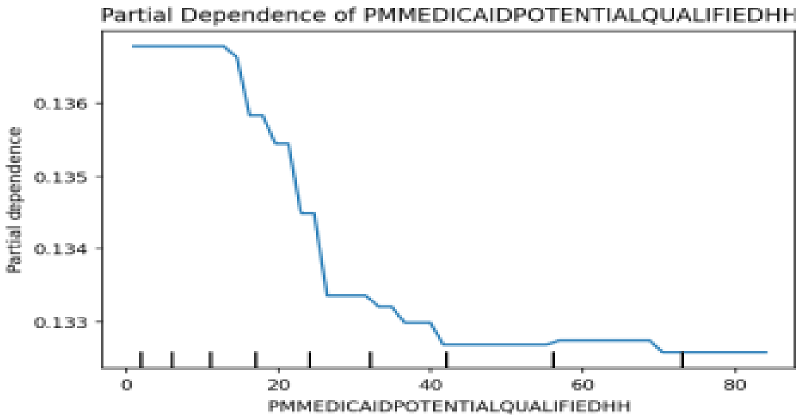


Figure 13. Partial Dependence Plot – Medicaid Qualified households

Value Chain Enthusiast (PM) is a proprietary vendor-derived ranking designed to identify households with strong price-conscious and value-oriented shopping behavior. Lower values on this scale indicate higher propensity for deliberate, affordability-focused purchasing. In the model, lower scores were associated with higher predicted response probabilities. This pattern is consistent with the broader observation from exploration analysis that lower-income and lower-value-score households were more responsive. These findings suggest that messaging emphasizing comparative value and budget-friendly options may resonate well with this segment.

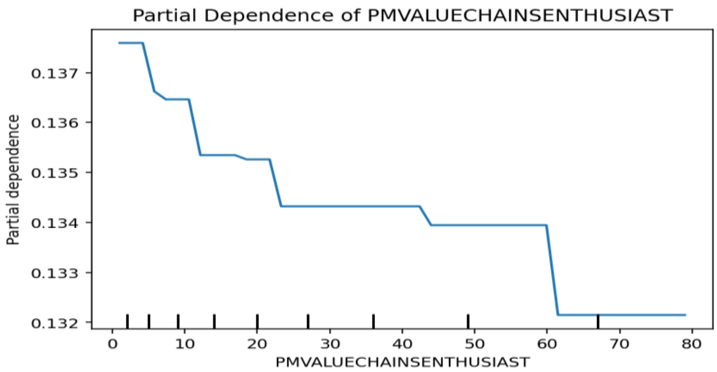


Figure 14. Partial Dependence Plot – Value Chain Enthusiast

Coin Collector (PM) and Medicare Plan D Prescription Purchaser (PM) emerged as less dominant predictors but likely act as proxies for age-related and retiree lifestyle segments. For example, coin collecting is a more prevalent hobby among older adults, while Medicare Part D participation directly signals senior status. These variables may capture latent behavioral or demographic dimensions not fully represented by other features, offering additional segmentation value despite lower standalone importance**.**

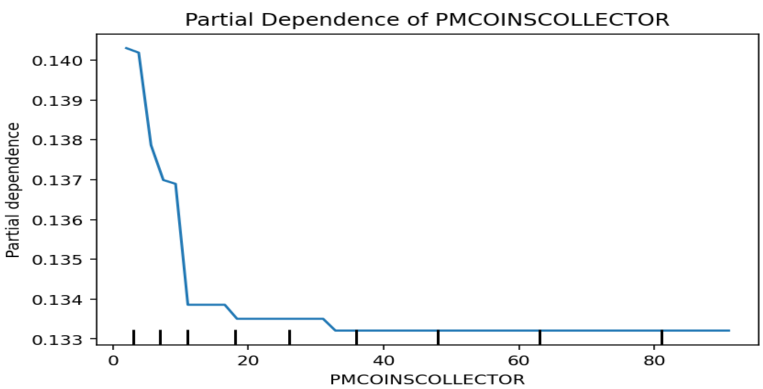


Figure 15. Partial Dependence Plot – Coin Collection

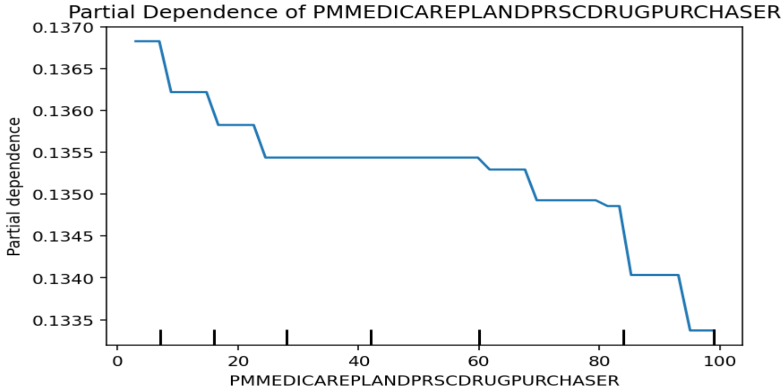


Figure 16. Partial Dependence Plot – Medicare Plan D Purchaser

Income Index and Mortgage (PM**)** served as important indicators of household financial health and stability. *Income Index*, which compares a household's income to county averages, demonstrated that lower-income households were more likely to respondconsistent with EDA findings showing elevated response rates among lower-income segments. Similarly, households with no mortgage obligation (renters or fully paid homeowners) showed higher predicted response. These patterns reinforce earlier exploratory insights about price-sensitive, less financially encumbered consumers being key targets for burial insurance offers**.**

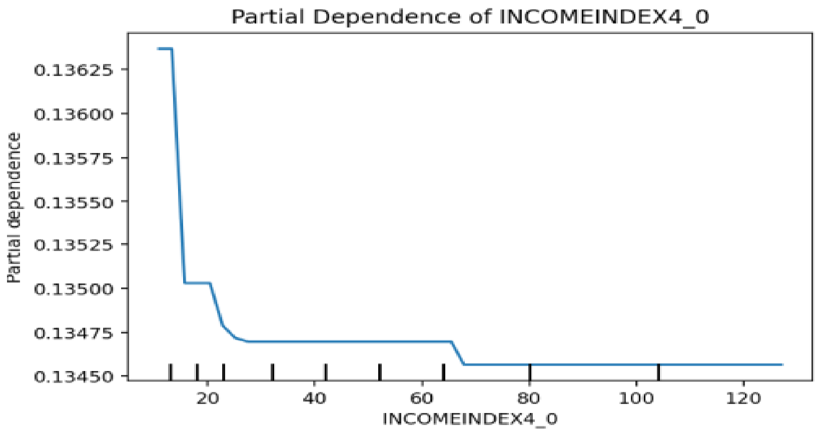


Figure 17. Partial Dependence Plot – Income Index

4. Discussions

4.1. Interpretation of Findings

This study demonstrates that predictive modeling can meaningfully improve target selection in burial insurance marketing by uncovering behavioral and socioeconomic signals of consumer responsiveness. The model's strongest predictors aligned closely with exploratory data analysis results, highlighting that financially constrained, mobile, and renter households are systematically more likely to respond.

Notably, variables such as Mobility Score, Medicaid Qualification, and Low Mortgage Status captured key aspects of economic vulnerability and housing stability, supporting the interpretation that these segments are highly sensitive to funeral cost planning. Behavioral indicators such as Sweepstakes Participation and Value-Conscious Purchasing (e.g., Value Chain Enthusiast) further suggest that consumers drawn to promotional offers and budget-focused shopping may be more receptive to burial insurance's promise of affordability and peace of mind. The use of Historical DMA Response Rates also reinforces the role of localized marketing knowledge in optimizing outreach strategies, showing that geographic differences can reveal distinct market receptivity profiles.

4.2. Theoretical and Practical Implications

These findings contribute to the segmentation, targeting, and positioning (STP) framework in marketing by illustrating how machine learning can operationalize segmentation in a multidimensional, data-driven manner. Rather than segmenting by single demographic variables, the model integrates a range of demographic, behavioral, and geographic indicators to generate individualized propensity scores. This enables marketers to prioritize outreach with much greater precision and relevance.

Additionally, the results support existing marketing theory around behavioral targeting and price-sensitive consumer decision-making, reinforcing that affordability salience, trust-building, and clear messaging are critical in low-income, older segments. By linking socioeconomic status, mobility, and promotional responsiveness to insurance purchasing intent, this work highlights actionable segmentation dimensions that marketers can leverage in campaign planning.

4.3. Managerial Recommendations

The modeling framework offers direct, practical guidance for burial insurance marketers:

- **Audience Selection:** Prioritize older, financially constrained renters and mobile households showing responsive behaviors (e.g., sweepstakes engagement).
- **Messaging Strategy:** Emphasize affordability, dignity, peace of mind for loved ones, and value-based benefits tailored to cost-sensitive consumers.
- **Channel Optimization:** Leverage DMA-level insights to localize offers, balance direct mail with other compliant outreach channels, and maximize ROI.
- **Exclusion Tactics:** Avoid targeting high-income, premium-oriented consumers who consistently demonstrate low responsiveness.

By integrating these insights into campaign workflows, marketers can reduce waste, improve conversion rates, and strengthen customer relationships through more relevant, personalized outreach.

4.4. Ethical Considerations

Marketing burial insurance for older, lower-income consumers demands special ethical care. The same modeling power that improves efficiency can risk overly aggressive or inappropriate targeting if not carefully managed. By precisely identifying likely responders, this approach can help reduce irrelevant mailings that burden consumers with low intent, respecting consumer time and privacy.

However, marketers should avoid exploiting informational asymmetries or financial stress, particularly in vulnerable groups. The best ethical practices should include clear, transparent communication, suitability checks where feasible and avoid fear-based or misleading messaging. The modeling framework presented here can support more respectful marketing by aligning offers with genuine consumer needs.

4.5. Limitations and Future Research

This study has several practical limitations. While the dataset is large and representative, it is sampled from a vendor-compiled database, which may introduce subtle selection effects. Some variables are vendor-derived scores with limited methodological transparency, which can constrain interpretability. Model generalizability should also be tested across other insurers or regions, as local market dynamics may differ. Finally, real-world effectiveness depends on integration with creative, offer, and channel strategy. Future research should include live campaign testing, tailored creative experimentation, ongoing monitoring, and model adaptation over time.

5. Conclusion

This study demonstrates the value of predictive modeling in enhancing targeting strategies for burial insurance marketing. By applying advanced machine learning to a rich dataset, the model identified key behavioral and demographic predictors of response, achieving a 1.8x lift in the top decile and capturing 43% of responders in the top three deciles. These findings support more precise segmentation, improved outreach efficiency, and better campaign ROI.

Beyond its applied value, this research contributes to marketing theory by extending data-driven targeting approaches to an underserved, price-sensitive insurance segment. It highlights how predictive analytics can uncover latent behavioral patterns in niche markets often overlooked in mainstream literature.

Future work should incorporate a broader variable set, utilize the full production dataset, and implement longitudinal monitoring. With continued refinement and integration into strategic planning, this framework offers a scalable foundation for sustained marketing effectiveness.

References

[1] Kotler P, Keller KL. *Marketing Management*. 15th ed. Pearson; 2016.
[2] Kotler P , Keller, K. L ref 1. above
[3] Wedel M, Kamakura WA. *Market Segmentation: Conceptual and Methodological Foundations*. 2nd ed. Springer; 2000.
[4] Blattberg RC, Kim BD, Neslin SA. *Database Marketing: Analyzing and Managing Customers*. Springer; 2008.
[5] Mayers D, Smith CW. Contractual provisions, organizational structure, and conflict control in insurance markets. *J Bus*. 1981;54(3):407–434.
[6] Cummins JD, Doherty NA. The economics of insurance intermediaries. *J Risk Insur*. 2006;73(3):359–396.

[7] Hansotia B, Wang P. Analytical challenges in customer acquisition. *J Direct Mark*. 1997;11(2):7–19.
[8] Baker T. *Risk, Insurance, and the Social Construction of Responsibility*. University of Chicago Press; 2002.
[9] Stone B, Jacobs R. *Successful Direct Marketing Methods*. 8th ed. McGraw-Hill; 2008.
[10] Financial Conduct Authority (FCA). GC19/03: Guidance on fair treatment of vulnerable customers [Internet]. 2019. Available from: https://www.fca.org.uk/publication/guidance-consultation/gc19-03.pdf [Accessed 2025 Jul 17].
[11] Verbeke W, Martens D, Mues C, Baesens B. Building comprehensible customer churn prediction models with advanced rule induction techniques. *Expert Syst Appl*. 2011;38(3):2354–2364. doi:10.1016/j.eswa.2010.08.023.
[12] Tummalapalli V, Konakalla K. Statistical techniques for feature selection in machine learning models. *Int J Innov Res Multidiscip Pursuit Stud*. 2025;13(3):1–8. doi:10.37082/IJIRMPS.v13.i3.232566.
[13] Lemmens A, Croux C. Bagging and boosting classification trees to predict churn. *J Mark Res*. 2006;43(2):276–286. doi:10.1509/jmkr.43.2.276.
[14] Brownlee J. *Probability for Machine Learning: Discover How to Harness Uncertainty with Python*. San Francisco: Machine Learning Mastery; 2020.
[15] Thirunagalingam, A. (2024). AI-Powered Continuous Data Quality Improvement: Techniques, Benefits, and Case Studies. Benefits, and Case Studies (August 23, 2024).
[16] P. K. Maroju, "Leveraging Machine Learning for Customer Segmentation and Targeted Marketing in BFSI," International Transactions in Artificial Intelligence, vol. 7, no. 7, pp. 1-20, Nov. 2023
[17] L. N. R. Mudunuri, "Artificial Intelligence (AI) Powered Matchmaker: Finding Your Ideal Vendor Every Time," FMDB Transactions on Sustainable Intelligent Networks., vol.1, no.1, pp. 27–39, 2024.
[18] Sehrawat, S. K. (2023). The role of artificial intelligence in ERP automation: state-of-the-art and future directions. *Trans Latest Trends Artif Intell*, *4*(4).
[19] Sudheer Panyaram, (2023), AI-Powered Framework for Operational Risk Management in the Digital Transformation of Smart Enterprises.
[20] Bhagath Chandra Chowdari Marella, "Scalable Generative AI Solutions for Boosting Organizational Productivity and Fraud Management", International Journal of INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING, vol. 11, no.10, pp. 1013–1023, 2023.
[21] Mohanarajesh, Kommineni (2024). Develop New Techniques for Ensuring Fairness in Artificial Intelligence and ML Models to Promote Ethical and Unbiased Decision-Making. International Journal of Innovations in Applied Sciences and Engineering 10 (1):47-59.