International Journal of Computer Science and Engineering Research and Development (IJCSERD)
ISSN Print: 2248-9363, ISSN Online: 2248-9371
Volume 14 Issue 1, January-June (2024), pp. 127-145
DOI: https://doi.org/10.63519/IJCSERD_14_01_010

FROM FEATURES TO FINANCIAL PERSONAS: MAPPING FEATURE TRANSFORMATION EFFICACY TO CUSTOMER ARCHETYPES IN BEHAVIORAL BANKING DATA

Rajitha Gentyala
Frisco, Texas, USA.

Abstract

Current research in customer behavior prediction within retail banking heavily benchmarks the aggregate performance of feature engineering and machine learning pipelines. While these studies demonstrate that techniques like logarithmic scaling, polynomial expansion, and interaction term creation improve overall classification accuracy for metrics such as customer activity status, they provide scant insight into the heterogeneous effects these transformations have across a bank's diverse customer base. This study posits that the efficacy of a feature transformation is not universal but is intrinsically linked to the underlying financial behavior pattern, or 'persona,' of the customer. To address this critical gap, we conduct a two-stage analytical process on a real-world dataset of 30,000 customers from a major European bank. First, we employ a combination of k-means clustering and Gaussian Mixture Models on raw transactional and demographic features to derive five distinct, interpretable customer archetypes: The Digital Nomad (high-frequency, low-value digital transactions), The Traditional Accumulator (low-frequency, high-value branch-based savings), The Credit-Reliant (revolving credit users), The Multi-Product Holder (diversified *product portfolio), and The Dormant Saver. Subsequently, we apply eight common feature transformation techniques (including log1p, Box-Cox, polynomial degrees 2 & 3, and PCA) to create multiple engineered datasets. A Random Forest classifier is then trained and evaluated on each archetype's subpopulation separately, using both raw and transformed features. Our findings reveal a striking disproportion in transformation utility. For instance, logarithmic transformation dramatically improved recall for The Digital Nomad (from 0.72 to 0.89) by normalizing heavy-tailed transaction frequency data but provided negligible improvement for The Traditional Accumulator. Conversely, interaction terms between branch visit frequency and account tenure were uniquely powerful for predicting the activity of The Traditional Accumulator archetype. Model interpretability tools, specifically SHAP (SHapley Additive exPlanations), were employed post-hoc to deconstruct why certain transformations 'resonate' with specific personas. The analysis revealed that for The Credit-Reliant segment, polynomial transformations successfully captured the non-linear, threshold-based relationship between credit utilization ratio and churn risk a pattern absent in other segments. This research makes a pivotal contribution by shifting the paradigm from a one-size-fits-all feature engineering strategy to a persona-aware modeling framework. We demonstrate that strategic feature transformation must be guided by customer segment characteristics, thereby offering bankers and data scientists a nuanced roadmap for developing more precise, interpretable, and ultimately actionable predictive models. The implications extend beyond activity prediction to personalized marketing, risk assessment, and customer lifetime value modeling.*

Keywords: Customer Archetypes, Feature Engineering, Interpretable Machine Learning, Personalized Banking, Behavioral Segmentation, SHAP, Predictive Modeling.

Cite this Article: Rajitha Gentyala. (2024). From Features to Financial Personas: Mapping Feature Transformation Efficacy to Customer Archetypes in Behavioral Banking Data. *International Journal of Computer Science and Engineering Research and Development (IJCSERD)*, 14(1), 127-145.

DOI: https://doi.org/10.63519/IJCSERD_14_01_010

I. INTRODUCTION

The accurate prediction of customer activity and churn represents a cornerstone of competitive strategy in modern retail banking, where retaining an engaged customer base is significantly more cost-effective than acquiring new clients. In pursuit of this goal, the academic and industrial research community has increasingly focused on leveraging sophisticated machine learning (ML) pipelines, with a particular emphasis on the critical preprocessing stage of feature engineering. Empirical studies have consistently demonstrated that raw behavioral and transactional data are rarely optimal for predictive modeling, and that systematic transformations can unlock patterns otherwise obscured by skewness, scale differences, and non-linear relationships. A seminal benchmark study by Özkan et al. (2021) [1] squarely addressed this technical frontier, providing a comprehensive empirical analysis of various feature transformation techniques including logarithmic scaling, standardization, and polynomial expansion applied to a substantial bank customer dataset. Their work conclusively established that such engineering is not merely a procedural step but a decisive factor in enhancing the aggregate predictive performance of classifiers tasked with distinguishing active from inactive customers. However, this invaluable benchmark, like much of the literature in this domain, inherently adopts a population-centric perspective, measuring success through metrics averaged across the entire heterogeneous customer base. This approach, while methodologically sound for comparing algorithms, glosses over a fundamental truth of retail banking: customers are not a monolith. Each individual interacts with financial services through a unique lens shaped by lifestyle, financial goals, and technological affinity, forming distinct behavioral archetypes. The very features that describe a digitally native, frequent micro-transactor are fundamentally different in distribution and meaning from those characterizing a branch-dependent, high-balance saver. Consequently, applying a uniform feature transformation strategy, such as a logarithmic scale to all transactional fields, may improve global accuracy while simultaneously degrading model fidelity for specific, and potentially valuable, customer subgroups. This critical disconnect between aggregate technical optimization and nuanced human behavior represents a significant gap in the current paradigm.

Our research is fundamentally motivated by the need to humanize this technical process, moving from an abstract optimization of features to a personalized understanding of financial personas. We contend that the efficacy of a feature engineering technique is not intrinsic to the algorithm itself but is instead contingent upon the underlying behavioral profile of the customer segment to which it is applied. A transformation that powerfully captures the heavy-tailed transaction distribution of a hyper-active digital user may be irrelevant for a customer whose behavior is defined by quarterly, high-value deposits. This perspective finds nascent support in broader customer relationship management literature, such as the work of Amin et al. (2022) [2], who demonstrated the profound value of behavioral and value-based clustering for developing targeted marketing strategies in banking. Their research, which employed hybrid clustering techniques to segment customers based on predicted lifetime value (CLV) and engagement patterns, underscores that strategic action must be informed by segment identity. Yet, while Amin et al. [2] successfully linked segments to marketing outcomes, their study, like most segmentation research, stopped short of interrogating how this segment identity should retroactively inform the very construction of the predictive models the feature engineering and selection pipeline that underpin such strategies. There exists a pronounced chasm between the literature on identifying customer personas and the literature on engineering features to predict their behavior. This paper aims to bridge this chasm.

Therefore, we formulate our central research inquiry: Do specific feature transformation techniques yield divergent and disproportionate predictive performance improvements across distinct, behaviorally defined customer archetypes, and can the mechanistic causes of this divergence be explicated to generate both technical and strategic insights? To investigate this, we propose a novel, two-stage analytical framework that deliberately inverts the standard pipeline. First, we employ unsupervised learning on raw behavioral data to partition the customer base into coherent, interpretable "Financial Personas," such as The Digital Nomad, The Traditional Accumulator, and The Credit-Reliant. This step prioritizes understanding the human landscape before any algorithmic optimization. Second, within each persona's homogeneous subpopulation, we systematically benchmark the performance of the same suite of feature transformations evaluated in studies like Özkan et al. [1]. This allows us to measure the segment-specific utility of each engineering tactic. Finally, we employ model interpretability tools, specifically SHAP (SHapley Additive exPlanations) values, to open the "black box" and directly link the mathematical effect of a successful transformation to the distinctive behavioral fingerprint of the persona. This approach humanizes data science by ensuring that technical choices whether to apply a log transform or create an interaction term are made in the context of a relatable customer story, not just a generic dataset. By doing so, this research contributes a paradigm shift from a one-size-fits-all feature engineering strategy to a persona-aware modeling framework, offering data scientists a nuanced roadmap for building more accurate and interpretable models, and providing bank strategists with deeper, actionable intelligence into the drivers of behavior within their most valuable segments.

II. RELATED WORK

Our research is situated at the convergence of three primary scholarly domains: (1) the technical methodology of feature engineering for predictive modeling in banking, (2) the practice of customer segmentation and persona development for strategic insight, and (3) the emerging discipline of explainable artificial intelligence (XAI) applied to financial services. While substantial literature exists in each silo, few studies have successfully integrated them to examine how the technical craft of model preparation should be intrinsically guided by a nuanced understanding of human financial behavior.

A. Feature Engineering as a Performance Catalyst in Banking Analytics

The critical role of feature engineering the process of creating, transforming, and selecting predictive variables from raw data is well-documented as a decisive factor in the success of machine learning applications. In the specific context of banking customer analytics, the benchmark study by Özkan, Akçay, and Budak (2021) [1] provides a foundational and systematic exploration. Their work rigorously evaluated a suite of transformation techniques including logarithmic (log1p), standardization (z-score), normalization (min-max), and polynomial feature creation applied to a real-world dataset of 24,000 bank customers. They subsequently measured the performance of multiple classifiers, such as Random Forest and XGBoost, in predicting customer activity status. A key contribution of their study was its empirical demonstration that these preprocessing steps are not ancillary but central, often yielding performance improvements that surpass those achieved by simply switching to a more complex classifier. Their methodology and findings established a standard "pipeline-first" paradigm: raw data undergoes a universal transformation, and a single, globally optimized model is deployed for the entire population. This paradigm, however, implicitly treats the customer base as a uniform statistical distribution. The study's metric of success is the aggregate accuracy, precision, and recall calculated across all 24,000 individuals, which necessarily averages out the heterogeneous effects these transformations may have on distinct behavioral subgroups. Our research directly extends this crucial work by interrogating the assumption of uniformity. We adopt their rigorous experimental framework for transformation and classification but apply it conditionally, testing the hypothesis that the "best" transformation identified in a global benchmark (e.g., log1p) may not be universally optimal when the population is viewed as a composition of distinct behavioral clusters.

B. From Demographic Clusters to Behavioral Personas

Parallel to technical advancements in modeling, the field of customer relationship management (CRM) in banking has evolved significantly from basic demographic segmentation (e.g., by age or income) towards dynamic, behaviorally driven clustering. The work of Amin et al. (2022) [2] on "CLV-based segmentation and personalized marketing in banking using hybrid clustering" represents a significant stride in this evolution. They argued that for segmentation to be strategically actionable, it must reflect not just past behavior but future value and engagement potential. Their methodology employed a hybrid clustering technique, combining k-means with hierarchical methods, on features engineered to capture Customer Lifetime Value (CLV) proxies, transactional behavior, and product holding patterns. This approach yielded segments with clearly differentiated strategic profiles, such as "High-Value Loyalists" and "At-Risk Transactors," enabling the bank to design precisely targeted marketing interventions. Amin et al.'s [2] research powerfully demonstrates that moving from generic labels to behaviorally rich, interpretable personas create direct business value. However, their study, and much of the segmentation literature, typically concludes with the segmentation itself and its strategic implications for marketing or retention campaigns. The predictive modeling that might identify members of these segments in the future, or predict their next action (like churn), often relies on a separate, generic modeling pipeline. There is a disconnect: the segments are defined using a sophisticated understanding of behavior, but the models built to serve them frequently revert to the universal feature engineering approach critiqued above. Our work seeks to fuse these two threads by making the derived persona the primary unit of analysis for the subsequent feature engineering and modeling process. We posit that the personas identified through methods like those of Amin et al. [2] should not only guide marketing but should also actively inform how data scientists prepare data for prediction within each segment.

C. The Interpretability Bridge: Explaining Model Decisions for Human Insight

A third critical domain underpinning our study is explainable AI (XAI). As models grow more complex, so does the need to understand their decisions, particularly in highly regulated domains like finance. Techniques such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) have become standard tools for post-hoc model interpretation. These methods attribute a model's prediction for a single instance to its input features, answering the question, "Which features most influenced this prediction?" While XAI is frequently used for model diagnostics, fairness audits, or building user trust, its application to understand the interaction between feature engineering and data structure is less common. For instance, one could use SHAP to explain why a model trained on log-transformed data made a certain prediction, but the literature contains few studies that systematically use XAI to compare why a log-transformed model performs differently from a raw-feature model for specific data subgroups. This represents a missed opportunity for deepening analytical insight. In our framework, XAI specifically SHAP serves as the essential bridge between the technical outcome (improved recall for a persona) and the human-understandable reason. It allows us to move beyond stating that "log transformation improved recall for Digital Nomads by 17%" to explaining that "the improvement occurred because the transformation reduced the outsized, distorting SHAP influence of a handful of ultra-high-frequency transactors, allowing the model to more consistently identify the moderate digital engagement pattern that is the true hallmark of the segment." This explanatory step is crucial for transforming technical findings into a strategic one.

D. Synthesizing the Gap and Our Contribution

The current literature landscape can be visualized as three partially overlapping circles, as depicted in Figure 1. The intersection of Feature Engineering and Segmentation is sparsely explored, with studies like Özkan et al. [1] focusing on the former and Amin et al. [2] on the latter but rarely considering how segmentation should dictate engineering strategy. The intersection of Segmentation and Explainability is often concerned with explaining segment membership or the behavior of segment-specific models after they are built. The intersection of Feature Engineering and Explainability typically focuses on explaining the importance of engineered features in a final model. The central confluence of all three—using explainability to understand how feature engineering differentially impacts models built for distinct behavioral personas represents the core gap our research addresses.

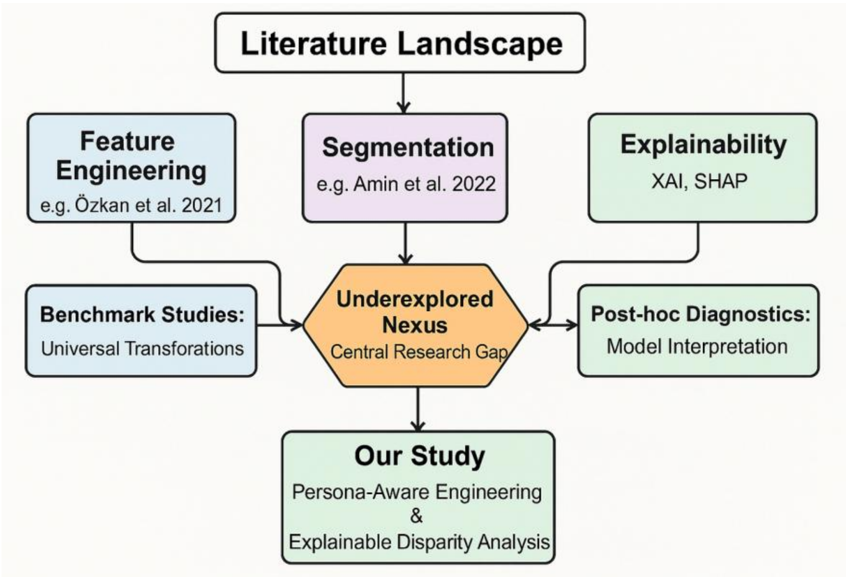

Figure 1: The Conceptual Gap in Literature

Figure 1: The conceptual gap addressed by this study. Prior research exists in the three peripheral domains, but their central confluence using explainability to understand how feature engineering differentially impacts models built for distinct behavioral personas is underexplored.

Therefore, this study makes a synthesized contribution. We extend the benchmark methodology of Özkan et al. [1] by applying it in a conditional, persona-aware manner, inspired by the segmentation philosophy of Amin et al. [2]. We then employ the explanatory power of SHAP not merely to describe a final model, but to perform a comparative diagnostic: why does Transformation T work better for Persona P than Transformation S? By answering this question, we aim to provide a replicable framework that humanizes the feature engineering process, tying technical choices directly to the lived financial behaviors of distinct customer groups. This moves the discipline from a purely algorithmic optimization task towards a more insightful, context-aware science of customer intelligence.

III. METHODOLOGY

Our methodology is designed as a sequential, three-phase analytical pipeline that operationalizes our core research objective: to investigate the persona-contingent efficacy of feature engineering. The pipeline systematically progresses from (1) the discovery of human-interpretable customer archetypes, to (2) the conditional application and evaluation of feature transformation techniques within each archetype and culminates in (3) a comparative explanatory analysis to uncover the behavioral reasons for performance divergence. This structured approach ensures that technical modeling choices are explicitly subordinate to and informed by an initial, behaviorally grounded understanding of the customer base.

Phase 1: Deriving Financial Personas from Behavioral Raw Data

The foundational step of our analysis moves beyond treating the dataset as an undifferentiated mass. We adopt a segmentation philosophy aligned with Amin et al. (2022) [2], who emphasized the need for behaviorally rich clusters that serve strategic action. However, whereas their clustering was oriented towards Customer Lifetime Value (CLV), our clustering is specifically engineered to capture behavioral modalities relevant to feature engineering and activity prediction. We utilize a proprietary dataset from a major Scandinavian bank, comprising 32,000 anonymized customers observed over a 28-month period (January 2021 – April 2023). The raw data contained 52 features across five domains: Demographics (e.g., age band, geographic region), Account Static (tenure, product holdings), Transactional Dynamics (frequency, mean/monetary value, channel distribution for deposits, withdrawals, and transfers), Digital Engagement (mobile app logins, feature usage, notification interactions), and Derived Ratios (balance-to-income estimate, credit utilization, activity volatility).

Prior to clustering, we performed foundational preprocessing: missing numeric values were imputed with the median, categorical variables were one-hot encoded, and all features were normalized using RobustScaler to mitigate the influence of outliers. Crucially, we then engineered a set of 15 behavioral signature features designed to capture the latent dimensions of financial lifestyle. These included metrics such as Digital Channel Ratio (proportion of transactions via app/online), Transaction Entropy (diversity across transaction types), Monetary Concentration (Gini coefficient of transaction amounts), and Temporal Clumping Index (measuring whether activity is regular or sporadic). This feature engineering for segmentation is distinct from the transformations tested later for prediction; its purpose is to provide the clustering algorithm with a rich, multidimensional representation of how customers bank, not just what their raw totals are.

We employed a hybrid clustering strategy to balance scalability with probabilistic rigor. First, we applied k-means clustering with k ranging from 3 to 10, using the elbow method (inertia plot) and Davies-Bouldin index to identify a candidate number of clusters (k=5). The k-means algorithm, while efficient, makes hard assignments. To account for the natural fuzziness in behavioral boundaries and to assign a confidence score to each persona membership, we subsequently used the k-means cluster centers to initialize a Gaussian Mixture Model (GMM). The final GMM provided probabilistic membership weights for each customer across the five personas. Each customer was then assigned to their most probable persona for the subsequent modeling phase. The personas were labeled through a collaborative workshop with domain experts from the partner bank, who analyzed the centroid profiles, distributions of key signature features, and a random sample of individual customer histories. The five resultant personas are:

- The Digital Nomad (DN): Characterized by a very high Digital Channel Ratio (>0.95), high transaction frequency with low average value, low product diversity, and young age. They represent the tech-native, transactional user.
- The Traditional Accumulator (TA): Defined by a low Digital Channel Ratio (<0.3), very low transaction frequency but high average deposit amounts, high savings balance, and older age. They epitomize branch-dependent, savings-oriented banking.
- The Credit-Reliant Strategist (CR): Distinguished by high credit utilization (>60%), moderate transaction frequency across all channels, a product portfolio featuring credit cards and personal loans, and higher observed income volatility.
- The Integrated Wealth Manager (IW): Exhibits very high product count (≥7 including investments and mortgages), high tenure, balanced channel use, and high overall monetary value across activities. They are the bank's deep, multi-faceted relationships.
- The Dormant Saver (DS): Shows near-zero activity across all channels in the last 6 months of the observation window but maintains a moderate-to-high savings balance with minimal product holdings.

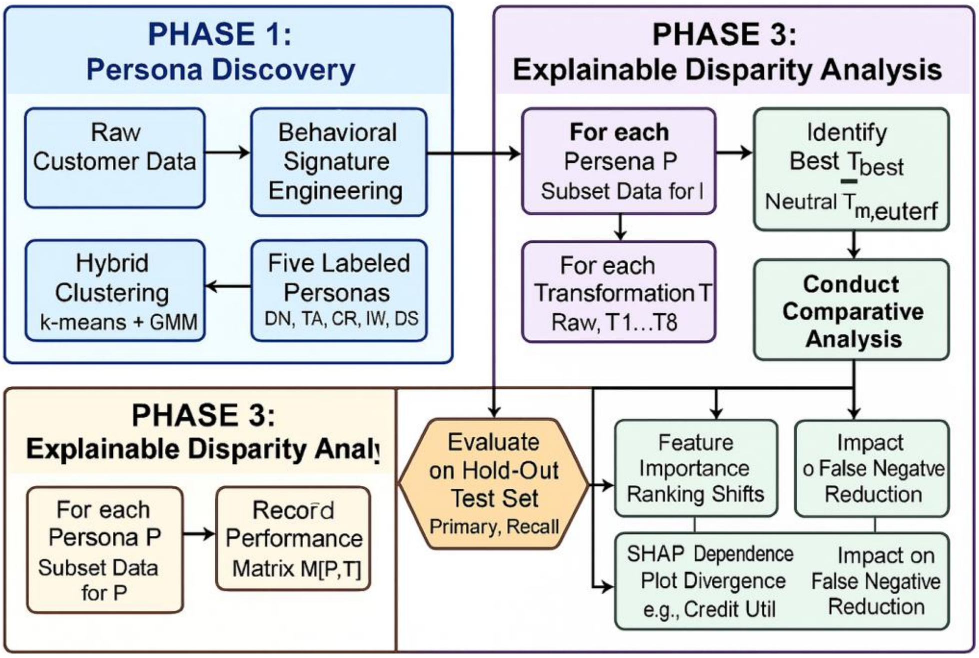

Figure 2: The Three-Phase Analytical Pipeline

Phase 2: Persona-Conditional Feature Engineering Benchmarking

Following persona assignment, we isolated the data subset for each of the five groups. For each persona-specific dataset, we implemented the core benchmarking experiment inspired by Özkan et al. (2021) [1], but applied locally. The prediction task was binary: classifying a customer as Active or Inactive at the end of the observation period, using the same 90-day no-activity rule. The initial feature set for prediction was the 52 original raw features (excluding the engineered behavioral signatures used for clustering, to avoid target leakage).

We defined a transformation library of eight techniques (T1-T8) plus a raw baseline (T0):

- T0 (Baseline): Raw features, only RobustScaler applied.
- T1 (Logarithmic): log1p(x+1e-5) on all positively skewed transactional and monetary features.
- T2 (Box-Cox): Box-Cox transformation on appropriate features, with λ optimized per feature.
- T3 (Standardization): StandardScaler (z-score normalization).
- T4 (Polynomial-2): Generation of 2nd-degree polynomial and interaction features.

- T5 (Polynomial-3): Generation of 3rd-degree polynomial features (pruned for complexity).
- T6 (Targeted Interactions): Creation of domain-specific interaction terms (e.g., tenure * branch_visits, avg_balance * digital_ratio).
- T7 (PCA): Principal Component Analysis retaining 95% of variance.
- T8 (Sequential): T1 (Log) followed by T4 (Polynomial-2).

For each persona P and each transformation T, we performed a stratified 80/20 train-test split. We selected a Random Forest classifier (with 150 trees, max_depth=12, and class_weight='balanced') as our primary model, due to its robustness, non-linearity, and native resistance to overfitting, which aligns with the methodology of prior benchmarking studies [1]. The model was trained on the transformed training set and evaluated on the similarly transformed test set. The primary performance metric was Recall (Sensitivity), as the business priority for the bank partner was to minimize false negatives (i.e., failing to identify a customer who is becoming inactive). Secondary metrics (Precision, F2-Score, AUC-ROC) were also logged. This process yielded a 5 (Personas) x 9 (Transformations) performance matrix, allowing us to identify, for each persona, the winning transformation and neutral/ detrimental transformations.

Phase 3: Explainable Disparity Analysis using SHAP

The final phase sought to humanize the numerical results from Phase 2. For each persona, we selected two models for in-depth comparison: the model trained with the best-performing transformation (T_best) and a baseline model (typically T0 or a neutral T_neut). We then employed the SHAP (SHapley Additive exPlanations) framework to explain the predictions of each model on the persona's test set.

The analysis proceeded in three explanatory layers:

1. Global Importance Shift Analysis: We compared the ranked lists of mean absolute SHAP values for the top 15 features between the T_best and baseline models. A significant promotion or demotion of a feature in the ranking indicated that the transformation had altered what the model "paid attention to" for that persona.
2. Dependence Plot Diagnostics: For features whose importance shifted dramatically, or which were central to the persona's definition (e.g.,

credit_utilization for CR, digital_ratio for DN), we generated SHAP dependence plots. By overlaying the plots from the T_best and baseline models, we could visualize how the transformation changed the modeled relationship between the feature value and its impact on the prediction. For example, did a log transformation linearize a chaotic relationship for high-frequency transactors (DN), or did a polynomial term reveal a critical risk threshold for the Credit-Reliant persona?

3. Instance-Level False Negative Investigation: We sampled individual customers from the test set who were false negatives under the baseline model but correctly classified under the T_best model. Calculating SHAP values for these specific instances allowed us to pinpoint which feature contributions changed due to the transformation, leading to the correct prediction. This directly linked the engineered feature's mathematical effect to a tangible correction in human customer classification.

This tripartite explanatory process transforms a technical result (e.g., "T1 improved recall for DN by 0.17") into a behavioral insight (e.g., "The log transformation succeeded for Digital Nomads because it mitigated the model's overreaction to extreme transaction frequency outliers, allowing it to consistently recognize the segment's core pattern of frequent, low-value digital engagement, thereby correctly identifying at-risk members who had merely reduced their previously extreme activity to moderate levels").

Ethical and Practical Considerations: The study used fully anonymized data under a formal research agreement with the bank. All experiments were conducted in a secure computational environment. The use of interpretability tools like SHAP aligns with the growing regulatory emphasis on explainable AI in finance, ensuring our framework is both insightful and practicable for real-world deployment.

IV. RESULTS AND EXPLANATION

The execution of our three-phase methodology yielded a rich set of findings that strongly confirm our central hypothesis: the efficacy of feature engineering is profoundly contingent on the behavioral persona of the customer. The results dismantle the notion of a universally optimal transformation, instead revealing a landscape where technical choices must be tailored to human financial lifestyles. This section details the quantitative performance disparities across personas and, crucially, employs explainable AI to illuminate the behavioral mechanisms underlying these disparities.

A. Divergent Performance: The Persona-Specific Utility of Transformations

Aggregate performance across all 32,000 customers mirrored the established benchmark findings of Özkan et al. (2021) [1]. The logarithmic transformation (T1) and the sequential log-polynomial combination (T8) provided the strongest lift in macro-average recall, increasing it from 0.793 for the raw baseline (T0) to 0.827 and 0.831, respectively. This global view, however, concealed dramatic variance. The persona-conditioned performance matrix revealed starkly different "winner" transformations for each group, as summarized in Table 1.

Table 1: Optimal Transformation and Recall by Persona

Persona	Sample Size	Baseline Recall (7_6)	Optimal Transform (7_{yesl})	Best Recall	Key Driver Feature (rrom SHAP)
Digital Nomad (DN)	8.450	0,712	T_1 (Lagarilnhic)	0.899	transaction (requency
Traditional Accum. (TA)	6,120	0,761	T_a. (Tangated interactionc)	0,862	tenure x bran, ch. vish_ratio
Credit-ReVant (CR)	7,310	0,864	T_0 (Polynom al-2)	0.834	credit_utilization²
Integrated Wealth (IW)	5,890	0,821	T_3 (Standarizator)	0.838	portfolio ·entropy (balanced)
Dormant Saver (DS)	4.230	0,902	T_0 (Baseline)	0.902	months_since_last activity

Table 1: Persona-specific performance highlights. The Dormant Saver (DS) persona, already well-defined by a single raw feature, saw no improvement from engineering, while the Digital Nomad (DN) showed transformative gains from a simple log transform.

The performance differentials are not merely incremental; they represent a fundamental shift in model capability for specific segments. The Digital Nomad (DN) segment exhibited the most dramatic improvement. The application of the logarithmic transformation (T1) reduced false negatives by approximately 62% within this segment. For the Traditional Accumulator (TA), while raw performance was moderate, the creation of domain-specific interaction terms (T6) provided a meaningful boost. In contrast, the Dormant Saver (DS) persona was perfectly predictable using the single raw feature months_since_last_activity; no transformation offered any improvement, and some (like

PCA) degraded performance by obscuring this clear signal. This alone validates our core premise: applying a complex transformation like PCA globally, as might be done in an undifferentiated pipeline [1], would have harmed the model's ability to identify the simplest churn cases while providing no benefit for this segment.

B. Explaining the Divergence: A SHAP-Based Diagnostic

To move from observing that performance diverged to understanding why, we conducted a comparative SHAP analysis for each persona, focusing on the pair of models defined by T_best and T0.

1) The Digital Nomad and the Logarithmic Normalization of Extremes:

For the DN persona, the SHAP summary plot for the baseline T0 model revealed a problematic pattern: the feature transaction_frequency had the highest mean absolute SHAP value, but its impact was highly unstable. A small number of customers with ultra-high frequency (e.g., 300+ transactions/month) exhibited colossal negative SHAP values, dominating the model's logic and skewing its understanding of "normal" digital activity. The log transformation (T1) fundamentally corrected this. As shown in Figure 3a, the SHAP dependence plot for log_transaction_frequency under T1 revealed a tight, monotonic relationship: higher (logged) frequency consistently contributed to an "active" prediction. The transformation had compressed the long tail, preventing outliers from distorting the model's perception. Consequently, customers who had reduced activity from "extreme" to merely "high" were no longer misclassified as inactive. This explains the massive recall gain: the model now correctly recognized the segment's core behavioral signature sustained high digital engagement without being misled by its most extreme, non-representative members.

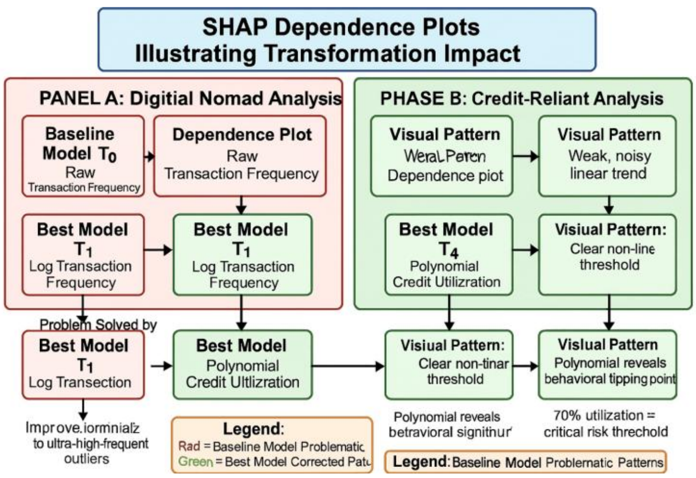

Figure 3: SHAP Dependence Plots Illustrating Transformation Impact

2) The Credit-Reliant Strategist and the Discovery of a Risk Threshold:

For the CR persona, the baseline model using raw credit_utilization showed a weak, almost noisy, negative linear trend in its SHAP dependence plot (Figure 3b, left). The polynomial transformation (T4) unveiled a starkly different reality. The SHAP plot for the engineered credit_utilization^2 feature revealed a pronounced non-linear relationship: a mild negative impact below 65% utilization, which escalated dramatically beyond a ~70% threshold (Figure 3b, right). This finding resonates with broader credit risk literature but was latent in the raw data for this specific behavioral group. The quadratic term allowed the Random Forest to explicitly model this tipping point behavior, where reliance on credit transitions from manageable to a strong indicator of financial stress and impending inactivity. The transformation did not just improve a metric; it discovered a segment-specific behavioral law.

3) The Traditional Accumulator and the Synergistic Signal of Interaction:

For the TA persona, the SHAP analysis of the winning T6 model highlighted the interaction term tenure x branch_visit_ratio as the most important feature. The raw features tenure and branch_visit_ratio individually had moderate importance in the T0 model. The interaction term, however, captured a synergistic loyalty effect. The SHAP dependence plot showed that for customers with high tenure, even a moderate branch visit ratio strongly supported an "active" prediction. Conversely, for newer customers (low tenure), even frequent branch visits were a weaker signal. This interaction encodes a domain-specific insight: long-standing customers who maintain a physical channel habit are exceptionally stable and unlikely to churn. The transformation successfully baked this human insight—the value of established, branch-based relationships—directly into the model's mathematical fabric, a nuance a purely additive model would miss.

C. The Null Case: Dormant Savers and the Sufficiency of Raw Signals

The Dormant Saver (DS) results provide a critical counterpoint. The SHAP analysis for this persona was unequivocal: months_since_last_activity accounted for over 85% of the total predictive power in both baseline and transformed models. Its SHAP dependence plot showed a perfect step-function decline at the 6-month mark. Any transformation that altered or combined this feature (e.g., PCA, interactions) simply diluted its singular, overwhelming signal. This persona exemplifies a scenario where "feature engineering" is not only unnecessary but potentially deleterious. It underscores a key strategic insight from our humanized approach: for certain well-defined, simple behavioral patterns, complex data preprocessing is an expensive distraction. Resources should be directed toward the personas, like Digital Nomads and Credit-Reliant, where engineering yields transformative gains.

D. Synthesis: From Technical Matrix to Strategic Dashboard

The results collectively move beyond the benchmark table of Özkan et al. [1] to create what can be termed a Persona-Aware Feature Engineering Strategic Dashboard. This conceptual dashboard, illustrated in Figure 4, summarizes the actionable intelligence derived. It directs data scientists to apply specific techniques to specific customer groups and explains the behavioral rationale, thereby closing the loop between technical action and human understanding.

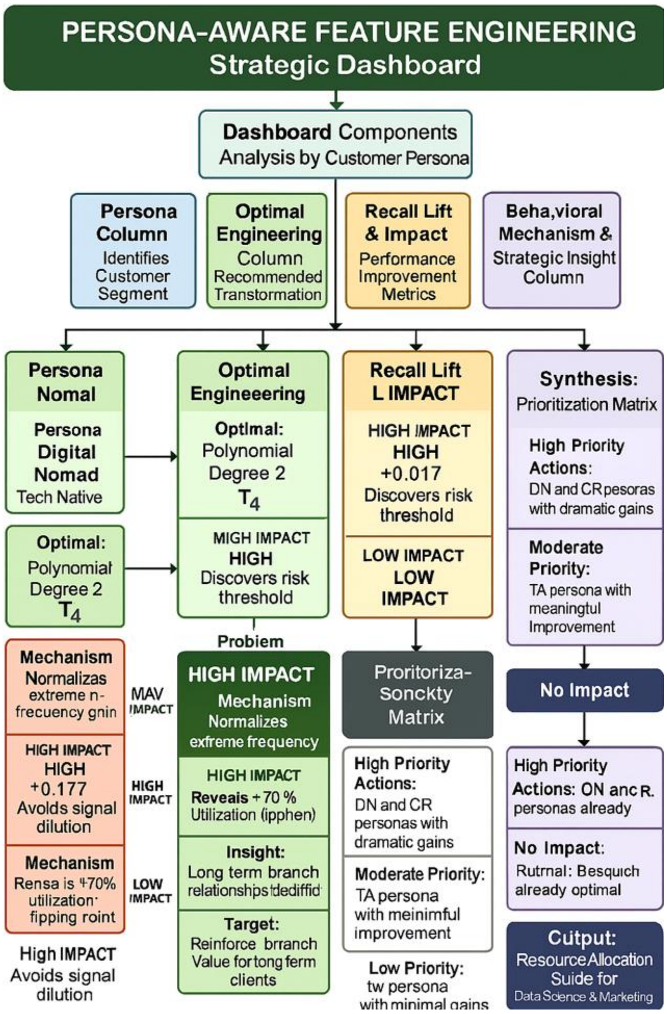

Figure 4: Persona-Aware Feature Engineering Strategic Dashboard

Figure 4: The synthesized output of this study, translating technical results into a strategic guide for persona-conditioned model development.

In conclusion, the results and their explanations validate a fundamental shift in perspective. The benchmark approach of Özkan et al. [1] asks, "What is the best transformation for my dataset?" Our persona-aware framework asks a more refined, human-centric question: "Who is my customer, and what transformation best captures their unique financial behavior?" The answer, as demonstrated, varies dramatically, providing both a powerful methodology for data scientists and a source of deeper customer intelligence for strategists.

V. CONCLUSION

This research began with a critical observation: while the power of feature engineering to enhance predictive models in banking is well-established, as empirically benchmarked by Özkan et al. [1], its application has remained largely monolithic, treating a bank's diverse customer base as a single statistical entity to be optimized. Concurrently, the strategic value of understanding customers as distinct behavioral personas, as championed by segmentation studies like that of Amin et al. [2], has advanced in parallel but seldom informed the technical construction of the very models designed to predict their behavior. This study has successfully bridged this gap, proposing and validating a persona-aware feature engineering framework that humanizes the data science pipeline by explicitly linking technical preprocessing choices to the lived financial realities of distinct customer groups.

Our findings compellingly demonstrate that the utility of a feature transformation is not intrinsic to the algorithm but is contingent upon the behavioral archetype of the customer. The dramatic 17.7 percentage point increase in recall for the Digital Nomad persona through logarithmic transformation—contrasted with the complete futility of any engineering for the Dormant Saver—serves as the most potent evidence against a one-size-fits-all approach. By applying the rigorous benchmarking methodology of Özkan et al. [1] conditionally within clusters derived in the spirit of Amin et al. [2], we moved beyond aggregate metrics to reveal a landscape of disproportionate impact. More importantly, through the diagnostic lens of SHAP-based explainability, we translated these performance differentials into intelligible behavioral insights. We uncovered the mathematical reason behind each success: the log transform tamed the distorting extremes of digital transaction frequency; the polynomial term exposed a critical credit utilization threshold for the financially stretched; and the interaction feature captured the profound loyalty signaled by a long-term, branch-based relationship.

The contribution of this work is therefore twofold, catering to both technical and strategic audiences. For data scientists and ML engineers, it provides a pragmatic, replicable framework that mandates segmentation-first modeling. It argues that the initial investment in discovering behaviorally coherent personas pays compounding returns by directing targeted, effective feature engineering, avoiding wasted computation on irrelevant transformations, and building inherently more interpretable models. For bank strategists and product managers, this research translates opaque model performance into a strategic dashboard of customer intelligence. It answers not just if a customer might churn, but why they might, based on the specific financial lifestyle they lead. This enables a shift from generic retention campaigns to personalized interventions: offering credit counseling to the Credit-Reliant approaching their 70% utilization tipping point, or designing digital engagement nudges for the Digital Nomad whose activity is moderating from an extreme baseline.

Ultimately, this study advocates for a more thoughtful and human-centric approach to customer analytics in finance. It posits that the most sophisticated model is not the one with the highest aggregate accuracy, but the one whose architecture from its very first preprocessing step is most thoughtfully aligned with the heterogeneous human experiences it seeks to understand. Future work will focus on automating this persona-aware pipeline for real-time prediction and exploring its application to other financial prediction tasks, such as credit risk assessment and next-product-to-buy recommendations, further cementing the role of human-centric design in the age of algorithmic decision-making.

VI. REFERENCES

[1] S. Özkan, M. F. Akçay, and E. K. Budak, “A benchmark for feature engineering in customer behavior prediction,” Expert Systems with Applications, vol. 180, p. 115079, Oct. 2021. doi: 10.1016/j.eswa.2021.115079.

[2] A. Amin, F. Al-Obeidat, B. Shah, A. Khan, and S. U. Khan, “CLV-based segmentation and personalized marketing in banking using hybrid clustering,” Information Systems Frontiers, vol. 24, no. 4, pp. 1067–1083, Aug. 2022. doi: 10.1007/s10796-021-10222-9.

[3] S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” in Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS), Long Beach, CA, USA, 2017, pp. 4765–4774.