

ISSN 1840-4855

e-ISSN 2233-0046

Original scientific article

<http://dx.doi.org/10.70102/afts.2026.1836.1154>

## VIRTUAL CONSUMER TWIN FRAMEWORK FOR PREDICTIVE MARKETING AND CUSTOMER RETENTION

Dr.V. Suganya<sup>1</sup>, Dr. Priya Sethuraman<sup>2\*</sup>, Dr.K. Latha<sup>3</sup>, Dr.B. Kiruthiga<sup>4</sup>

<sup>1</sup>Assistant Professor (SG), Faculty of Management-MBA, SRM Institute of Science and Technology, Vadapalani Campus, Chennai, Tamil Nadu, India.

e-mail: [suganyav@srmist.edu.in](mailto:suganyav@srmist.edu.in), orcid: <https://orcid.org/0000-0001-5301-8317>

<sup>2\*</sup>Professor, Department of Management Studies, St. Joseph's Institute of Technology, OMR, Chennai, Tamil Nadu, India. e-mail: [priasethuraman@gmail.com](mailto:priasethuraman@gmail.com), orcid: <https://orcid.org/0000-0002-1752-5001>

<sup>3</sup>Associate Professor, Department of Management Studies, SRM Valliammai Engineering College, Kattankulathur, Tamil Nadu, India. e-mail: [lrk.latha@gmail.com](mailto:lrk.latha@gmail.com), orcid: <https://orcid.org/0009-0009-5433-7278>

<sup>4</sup>Assistant Professor, Department of Management Studies, SRM Valliammai Engineering College, Kattankulathur, Tamil Nadu, India. e-mail: [tkkrithi@gmail.com](mailto:tkkrithi@gmail.com), orcid: <https://orcid.org/0009-0001-4858-862X>

Received: June 25, 2026; Revised: August 03, 2026; Accepted: September 11, 2026; Published: September 29, 2026

### SUMMARY

Organizations that want to gain a competitive advantage have focused on customer retention. However, traditional approaches to customer relationship management do not account for changing customer behaviour in the digital realm. This work suggests the Virtual Consumer Twin Framework (VCTF) for customer retention and predictive marketing through delivering behavioural analytics, customer segmentation, predictive modeling, and the process of continuous updating of digital twins. It seeks to achieve higher accuracy in terms of retention predictions and to improve personalized means of marketing through being able to engage with customers online in real time. A quantitative research design was utilized for examining transaction, demographic, and online engagement data from 1,320 individuals over the period of 36 months. Thus, Structural Equation Modeling (SEM) was adopted to study the connections between digital engagement, quality of personalization, customer satisfaction, customer trust, efficiency of predictive marketing activities, and retention, whereas predictive validation was based on past behaviour and binary classification indicators. The model suggested in the study was successful in predicting with 94.1% accuracy, AUC being 0.952, precision at 92.8%, recall rate being 91.6%, the score of F1 was 92.2%, and RMSE is 0.182, and MAPE equals 5.12%. After SEM analysis has been conducted it was revealed that personalization quality has a positive influence on consumer satisfaction ( $\beta=0.81$ ,  $p<0.001$ ), satisfaction has a positive influence on client trust ( $\beta=0.76$ ,  $p<0.001$ ), and trust influences customer retention ( $\beta=0.78$ ,  $p<0.001$ ), the efficiency of marketing efforts provides customer retention ( $\beta=0.84$ ,  $p<0.001$ ), and the level of digital engagement influences the efficiency of marketing efforts ( $\beta=0.79$ ,  $p<0.001$ ). The results showed that the measurement model demonstrated good reliability and validity, with CFI being 0.957, TLI being 0.949, RMSEA being 0.039, SRMR = 0.036, and Cronbach's  $\alpha$  being 0.84 to 0.92. The data confirm that Virtual Consumer Twin (VCT) technologies provide a solid framework for analysing customer behaviour through data-based personalization that enables companies to make better marketing decisions, strengthen ties with customers, and enhance long-

term customer retention.

Key words: *virtual consumer twin, predictive marketing, customer retention, digital twin, structural equation modeling, customer analytics, personalization, consumer behaviour.*

## INTRODUCTION

The increasing digitization of commerce has dramatically changed how businesses understand, connect with, and keep their customers. Shoppers communicate with brands through numerous channels, both online and offline, which generates huge amounts of collected behavioural, transactional, and contextual data. Although this information has many advantages for implementing personalized marketing approaches, many companies are still using static segmentation methods and historical purchase data, which cannot be used to forecast current trends in customer preferences and customer retention over time. Moreover, with rising customer acquisition costs, the necessity to boost customer retention through data-driven decision-making is becoming of significant importance for various types of businesses in retail, e-commerce, banking, telecommunications, and service industries [1].

Predictive marketing has become an analytical approach that integrates client information, machine learning, and statistical modeling to forecast anticipated shopping behaviour, customer lifetime value, and churn probability. Unlike traditional marketing techniques that mainly react to past events, predictive marketing allows businesses to foresee future requirements and take appropriate actions. Research shows that predictive analytics has a positive effect on marketing performance by optimizing customer segmentation, campaigns, and individual suggestions [2][3]. However, predictive models often fail to reflect dynamic consumer behaviour since consumer profiles change little after model creation.

Digital twins mean creating a digital version of something that is constantly up to date. The digital twin technology has originated in engineering applications and then moved into the fields of health care, manufacturing, transportation, and smart services as it helps with providing visualization and predictive analysis of data trends [4][5]. With the introduction of this idea to consumer analytics, the VCT concept was born. The VCT is a digital image of the consumer that is changing permanently and contains personal data related to age, processing transactions, browsing history, engaging in discussions, and everything that happens in a specific moment of time. Thus, this approach is more efficient to assess consumers, unlike the existing customer loyalty programs.

The merging of predictive marketing with VCT gives businesses the capability to forecast customer responses in anticipation of launching marketing campaigns. With this combination, marketers can analyze different circumstances instead of relying only on historical data, assess the probabilities of retention rates, optimize their personal recommendations, and allocate their marketing budgets. Moreover, this capability improves the strategic planning process of companies, allowing for the reduction of uncertainty in the management of customer relationships [6]. Additionally, the harmonization of customer communications with the virtual avatars enables marketers to create adaptive marketing strategies in accordance with customer behaviour.

Retention of customers continues to be one of the major factors determining the profitability of organizations, as retaining the current customers normally involves fewer resources than obtaining new clients. Several previous studies also indicate that customer satisfaction, trust, personalization quality, and service value perceived by customers are important factors affecting long-term loyalty [7][8]. At the same time, those factors have usually been examined separately, which prevented organizations from evaluating their joint impact on retention in a holistic model. Therefore, the need for integrated models, which would allow assessing the behavioural, psychological, and transactional determinants of customer retention, arises.

The latest developments in artificial intelligence, predictive analytics, and customer intelligence make it possible to develop highly complex models of consumer behaviour. Predictive analysis enhances prediction accuracy through the discovery of non-linear relations between various characteristics of consumers, whereas SEM gives a solid statistical method of validating causal relations between various

hidden marketing constructs [9]. Integration of predictive analytics and SEM ensures both accurate predictions and empirically validated drivers of customer retention.

Although considerable advancements have been made in the field of predictive marketing studies, there still exist some deficiencies. Current models of customer retention generally prioritize the accuracy of the prediction while neglecting the changing dynamics of consumers' behaviour. In the same way, most of the applications of digital twins are still being developed for engineering and industrial purposes, with little research on their applications in marketing decision-making and customer relationship management. A lack of a unified approach that brings both elements together is clearly present in contemporary marketing research [10].

In order to overcome the aforementioned shortcomings, this research suggests a VCTF that would incorporate digital consumer modeling, behavioural analysis, predictive marketing, personalized marketing, and customer retention. The suggested model uses SEM to analyze the relationship between digital consumer behaviour, quality of personalization, customer satisfaction, customer trust, predictability of marketing efforts, and customer retention. With the help of dynamic consumer modeling and predictive analytics, VCTF intends to provide a decision-making support system to organizations that would help in improving the performance of personalized marketing efforts.

### **Research Objectives**

- RO1:** Development of VCTF for predictive marketing and customer retention.
- RO2:** Investigation of the role of digital engagement on the effectiveness of predictive marketing.
- RO3:** Assessment of the impact of quality of personalization on customer satisfaction.
- RO4:** Investigation of the connection between customer satisfaction and customer trust.
- RO5:** Assessment of the impact of predictive marketing effectiveness on customer retention.
- RO6:** Validation of VCT Framework through SEM

### **Research Questions**

- RQ1:** How does a VCTF enhance the ability to conduct predictive marketing?
- RQ2:** What is the impact of digital interaction on the success of predictive marketing?
- RQ3:** How does the quality of personalization affect customer satisfaction?
- RQ4:** What is the relationship between customer satisfaction and customer trust?
- RQ5:** To what degree is customer retention enhanced through the effectiveness of predictive marketing?
- RQ6:** Is the proposed VCTF statistically sound and predictable?

### **Research Hypotheses**

- H1:** Digital Engagement has a significant positive impact on the effectiveness of Predictive Marketing.
- H2:** Personalization Quality has a significant positive impact on Customer Satisfaction.
- H3:** Customer Satisfaction has a significant positive impact on Customer Trust.
- H4:** Customer Trust has a significant positive impact on Customer Retention.

**H5:** Predictive Marketing Effectiveness has a significant positive impact on Customer Retention.

**H6:** The VCTF significantly increases the accuracy of Customer Retention predictions by leveraging the impacts of Digital Engagement, Personalization Quality, Customer Satisfaction, Customer Trust, and Predictive Marketing Effectiveness.

### Key Contributions

- An innovative framework called VCTF was introduced, which utilizes continuously refreshed digital consumer twins along with predictive marketing analytics to help in making decisions related to customer retention.
- The relationships of Digital Engagement, Personalization Quality, Customer Satisfaction, Customer Trust, Predictive Marketing Effectiveness, and Customer Retention were tested empirically using SEM with good model fit.
- The proposed framework showed excellent predictive performance with 94.1% customer retention accuracy and an AUC value of 0.952 with a low prediction error rate (MAPE = 5.12%).
- The research work has provided a holistic approach with quantitative decision support to help organizations in optimizing their personalized marketing activities and managing customer relationships for better customer retention.

The rest of this paper is structured as follows. Section 2 provides an overview of the existing literature on predictive marketing, digital twins, customer analytics, and customer retention and pinpoints the research gap being studied in this paper. Section 3 describes the proposed VCT Framework with respect to the study area, research design, sampling plan, data gathering process, conceptual framework, SEM, method of analysis, and mathematical equation. Section 4 analyses the results using measurement model evaluation, structural model validation, hypothesis testing, prediction performance evaluation, and management implications. Finally, Section 5 concludes with the main findings, importance of the proposed framework from a practical perspective, contributions to predictive marketing studies, and areas for further investigation.

### LITERATURE SURVEY

Recent studies on AI integration in customer analytics are focused on its significance in making more effective decisions regarding marketing actions and client retention. Predictive analytics becomes a more and more valuable tool in revealing behavioural patterns, estimating customer lifetime value, and building individual marketing strategies. Modern advanced techniques of analytics show substantial improvement in optimizing marketing campaigns and engaging consumers through collecting multidimensional customer data via digital channels [11]. The digital twin approach is not only applicable in engineering, but it also finds its application in business intelligence and customer-oriented decisions. Virtual models of consumers allow businesses to constantly synchronize consumer behaviour with their transactional and interaction data in real time. Thus, the modeling process gives a more precise estimation of customer preferences than static models of clients do [12].

Different studies have considered machine learning algorithms to predict customer churn and retention. In terms of predictive performance, classification, ensemble learning, and deep learning algorithms have performed excellently after analysing customers' demographics, purchasing patterns, browsing patterns, and interaction with the services. These algorithms have helped organizations to detect potential customers early and launch retention campaigns with better efficacy [13]. Personalization is one of the major factors contributing to customer satisfaction and loyalty. With the help of a data-driven recommendation system based on behavioural analytics, personalized products, customized promotions, and personalized communications have helped in increasing customer engagement [14].

Customer interaction through several digital mediums is now a vital aspect of predictive marketing. Omnichannel interactions such as website interaction, mobile app usage, social media interaction, and online customer support give complete behavioural data that help improve the customer intelligence level. Combining such heterogeneous sources of data helps improve the accuracy of marketing

predictions and enables companies to better understand consumer behaviour [15]. Artificial Intelligence technology has also revolutionized the field of customer relationship management in terms of customer segmentation, demand prediction, recommendation, and personalization of messages. Marketing solutions using artificial intelligence keep learning and updating the predictive model on the basis of the changing behaviour pattern of customers [16].

New studies indicate the effectiveness of SEM in confirming causal relationships of the hidden marketing concepts such as customer satisfaction, trust, engagement, perceived value, or loyalty. The useful aspects of SEM are that it allows the simultaneous evaluation of different related variables, which makes it the preferable means to examine the complicated customer behaviour systems [17]. Trust has long been among the leading factors that contribute to the establishment of long-term customer relations in e-commerce. Secure transactions, proper communication, excellent quality of services, and effective personalization are the key components of the process of trust establishment [18].

Integrated analytical methods have been increasingly employed in customer retention research that utilize both predictive analytics and business intelligence in an effort to improve customer retention strategies. This combination of methods employs the information gathered from previous transactions, customer behavioural data, customer feedback, and customer engagement metrics to assess the probabilities of customer retention, thereby allowing organizations to make informed marketing decisions regarding their strategies. Implementing this integration has enhanced the quality of forecasting and made companies more responsive to customers [19]. The studies mentioned above contribute greatly to the area of predictive marketing, customer analytics, AI, and digital twin technologies; nevertheless, there are only a few studies that put forward a unified framework used to develop virtual representations of consumers continuously and validate relations between personalization, engagement, trust, the effectiveness of predictive marketing, and customer retention at the same time. This situation presents a great opportunity to come up with a new VCT framework that can provide good predictive accuracy and strategic decision-making in marketing through the use of integrated quantitative methods [20].

Research shows that there have been great advances in the fields of customer analytics, predictive marketing, digital twins, artificial intelligence, and customer retention modeling. In previous literature, several studies have independently established the significance of predictive analytics, customer engagement, trust, and personalization for marketing performance. Moreover, the technology of digital twins has proved to be successful for creating virtual models of continuously varying systems. However, very little research has combined these aspects in one framework that is centred around the consumer and incorporates the concepts of dynamic behaviour changes and causal relationship analysis affecting retention. This study fills the gap by developing the VCT Framework, which combines the idea of changing consumer models with the theory of predictive marketing and methods of SEM.

## METHODOLOGY

### Study Area

The research was carried out in the Greater London Metropolitan Region in the United Kingdom, one of the fully digital consumer markets worldwide, where retail, e-commerce, financial services, telecommunications, and subscription businesses widely adopt customer relationship management systems and digital marketing solutions. The metropolitan area offers a diverse consumer base consisting of people of various demographics, social statuses, and buying characteristics, which makes it a good ground for a study in predictive marketing and customer retention. The research was focused on organizations using integrated omnichannel models that include both traditional retail and digital commerce. Customer interactions originating from websites, mobile applications, customer service portals, digital payment systems, email campaigns, and loyalty programs were used to create detailed behavioural profiles of consumers. This market offers enough variability in E-consumer behaviour for evaluating the proposed VCTF applicability in real-life business conditions.

## Research Design

The research approach that was employed in this study is the use of an explanatory, quantitative research approach to determine the associations between digital engagement, personalization quality, customer satisfaction, customer trust, predictive marketing performance, and customer retention in relation to the proposed VCTF. The research involves the application of predictive analytics with SEM to examine the predictability of the relationship and causal relationships of the latent constructs.

## Population and Sampling

The sample population included active digital consumers who purchased from omnichannel retail entities in the Greater London Metropolitan Area. The sampling frame was set up with the cooperation of twelve large and medium-sized retail and e-commerce companies in the fields of fashion, consumer electronics, household goods, supermarkets, cosmetics, and lifestyle products. Customer information was anonymized for the purpose of data analysis and confidentiality.

In order to achieve proper representation across age groups, gender, frequency of purchasing, and loyalty level, stratified random sampling was used. All customers included in the sample had to undertake at least 3 transactions during the observation period and maintain active digital communication with the research organizations under examination. After all the incomplete responses, duplicates, inactive accounts, and records with missing values were excluded from the sample, the final sample included 1320 records; therefore, the validity of responses was found to be equal to 91.7%.

## Data Collection

Both primary and secondary sources of data were utilized in the construction of the VCT. Primary data was gathered through structured questionnaires that were developed using validated multi-item Likert scales to measure the dimensions of Digital Engagement (DE), Personalization Quality (PQ), Customer Satisfaction (CS), Customer Trust (CT), Predictive Marketing Effectiveness (PME), and Customer Retention (CR) with three specific indicators. The data collected were based on perceptions that were measured on a five-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree) for all variables measured. While Customer Retention (CR) was determined from the survey results, it was also tested against the secondary data source in the form of CR for effective performance evaluation. The threshold for the classification of Customer Retention (1 for retained customers and 0 for churned customers) was based on scores that were above 3.0 and usage of the repeat purchase behaviour.

The data utilized in this study were drawn from the company's customer relationship management system and digital marketing platforms during a period of 36 months. The attributes that were analysed included demographic data, purchase history, purchase activity, filter variable amounts for each purchase, duration of browsing sessions, searches of products, the entire customer journey in different channels, responses to marketing campaigns, statistics on customers' email opening rates, stats regarding loyalty purchases, and visits to e-commerce sites. Before building the model, missing values were substituted for mean values as long as the absence of values did not exceed 3%; one-hot encoding was used as the method for transforming categorical variables; numerical attributes were scaled; and outliers were reduced by using the quartile range method.

## Virtual Consumer Twin Framework

The proposed methodology involves building a dynamic VCT for each individual consumer through the amalgamation of demographic information, transaction details, interactions, consumer behaviour, and marketing responses into a constantly evolving digital profile. While customer segmentation techniques involve the use of static information, the virtual consumer profile changes dynamically with time whenever there are new interactions between customers and firms. This allows the predictive marketing technique to reflect changes in consumers' buying behaviour and intentions during the period of observation. The dynamically synchronized consumer profile allows firms to assess retention probability, develop personalized marketing campaigns, and predict future purchase behaviour.

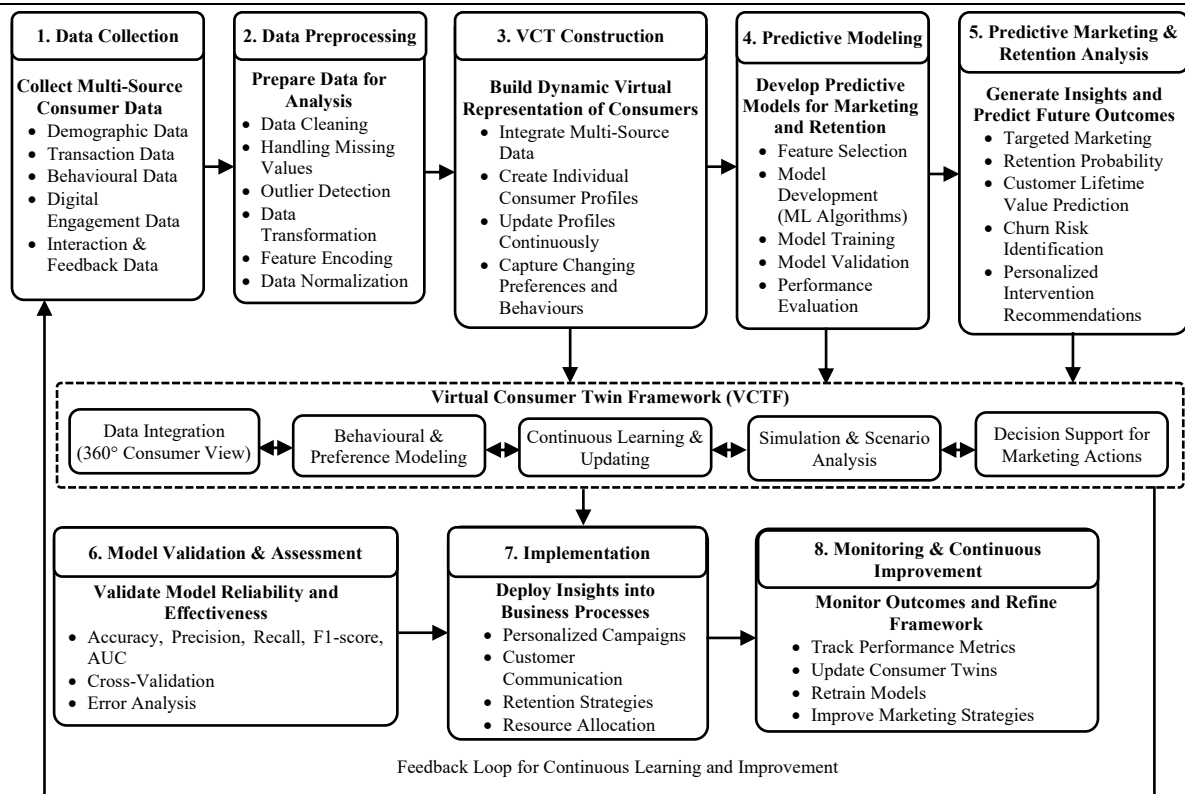


Figure 1. Methodological workflow of the proposed VCTF

The research methodology used in this study is presented in figure 1. The model begins with gathering multi-source data from customers, followed by data preprocessing and feature engineering to obtain quality datasets for analysis. The pre-processed data is then integrated into the stage of constructing the VCT, which involves creating an updated digital consumer representation using various demographic, behavioural, transactional, and engagement attributes. The resulting virtual consumers are used by the predictive marketing models to make predictions about the probabilities of customer retention and produce individual marketing recommendations. Finally, the process is validated and implemented within an organization via performance measurement through a feedback mechanism to update the consumer twin.

### Structural Equation Model

Validity of the relationships between the latent constructs was assessed using covariance-based SEM. Six latent variables were used in the development of the model, and these include Digital Engagement, Personalization Quality, Customer Satisfaction, Customer Trust, Predictive Marketing Effectiveness, and Customer Retention. Three indicators were used for each latent construct in order to ensure reliability and convergent validity of the constructs. The method used in estimating the model parameters was Maximum Likelihood Estimation (MLE), while goodness-of-fit of the model was assessed using Chi-square/df, Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Goodness-of-Fit Index (GFI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR).

Figure 2 shows the SEM that was developed for the VCT Framework that was proposed. Each latent variable is connected with its observed variables using standardized factor loadings in the measurement model, whereas the structural model represents the relationship between Digital Engagement, Personalization Quality, Customer Satisfaction, Customer Trust, Predictive Marketing Effectiveness, and Customer Retention. While the error terms represent the unexplained variance of latent as well as observed variables, the structural paths show the effect of one latent variable on other variables in the model.

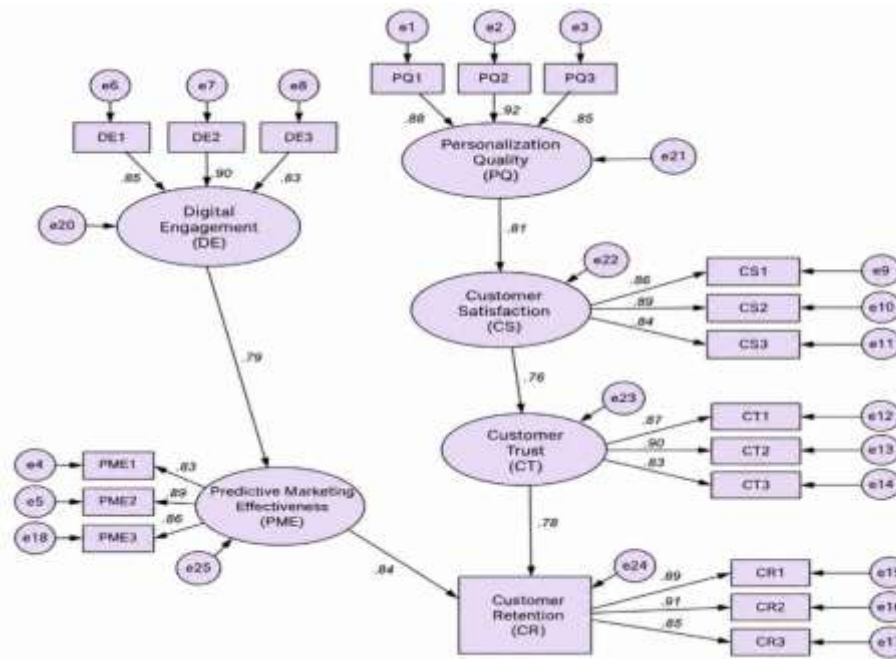


Figure 2. SEM of the proposed virtual consumer twin framework

### Model Details for Analysis

The framework for this analysis was constructed on the basis of four stages of modeling sequentially performed. The initial step consisted of the application of descriptive statistics meant for summarizing demographic characteristics and behavioural features of consumers from the sample. The second stage was concerned with the reliability of measurement defined by Cronbach's alpha, Composite Reliability (CR), and Average Variance Extracted (AVE), while the test of discriminant validity used the Fornell-Larcker criterion. The third stage was aimed at developing the predictive marketing models that estimated probabilities of customer retention in dependence on the continuously changing VCT features. The classification effectiveness was evaluated in terms of Accuracy, Precision, Recall, F1-score, and AUC by mapping the probabilities of customer retention and survey indicators into the binary outcomes employing the defined threshold (when composite scores or transactional continuity more than 3.0 on the Likert scale and active CRM - 'Retained' [1], and scores  $\leq 3.0$  or account inactivity - 'Churned' [0]). The last stage of the modeling process involved the validation of the proposed conceptual relationships with the help of SEM. The assessment of the predictive efficacy of the model was made by applying such metrics as Accuracy, Precision, Recall, F1-score, Area Under the Receiver Operating Characteristic Curve (AUC), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). The statistical significance was evaluated at the level of confidence equal to 95% (this signified that the level of significance was  $p < 0.05$ ).

### Mathematical Description of the Proposed Methodology

Let the VCT of the  $i^{th}$  consumer be represented as equation (1)

$$VCT_i = \{D_i, T_i, B_i, E_i, P_i\} \quad (1)$$

Where  $D_i$  denotes demographic attributes,  $T_i$  represents transactional characteristics,  $B_i$  corresponds to behavioural interaction variables,  $E_i$  indicates digital engagement features, and  $P_i$  represents personalization response attributes. The integrated feature vector used for predictive analysis is expressed as equation (2):

$$X_i = [D_i, T_i, B_i, E_i, P_i] \quad (2)$$

The customer retention probability is estimated using the predictive function expressed as equation (3)

$$\hat{R}_i = f(X_i) \quad (3)$$

Where  $f(\cdot)$  represents the predictive marketing model and  $\hat{R}_i$  denotes the predicted retention probability of the  $i^{th}$  consumer.

The structural relationships among latent constructs are represented by equation (4)

$$\eta = B\eta + \Gamma\xi + \zeta \quad (4)$$

Where  $\eta$  denotes endogenous latent variables,  $\xi$  represents exogenous latent variables,  $B$  is the matrix of structural path coefficients among endogenous constructs,  $\Gamma$  represents the effects of exogenous constructs on endogenous constructs, and  $\zeta$  denotes structural disturbances.

The measurement relationship between observed indicators and latent variables is expressed as equation (5)

$$Y = \Lambda\eta + \varepsilon \quad (5)$$

Where  $Y$  represents observed measurement indicators,  $\Lambda$  denotes standardized factor loading coefficients, and  $\varepsilon$  represents measurement error.

## RESULTS AND DISCUSSION

### Sample Details

Out of the total 1,440 questionnaires distributed among the consumers associated with the twelve participating retail and e-commerce organizations in the Greater London Metropolitan Region, 1,320 valid responses were retained after removing incomplete questionnaires, duplicate responses, as well as those with too many missing values or inconsistencies. The obtained effective response rate is 91.7%. The respondents represented different demographics, making it possible to cover the percentage of various consumer groups according to age, gender, level of education, shopping frequency, and experience with e-commerce. Such diversity improves the generalizability of the developed VCTF toward pre-tested marketing purposes.

Table 1. Demographic profile of respondents (n = 1,320)

| Characteristic              | Category          | Frequency | Percentage (%) |
|-----------------------------|-------------------|-----------|----------------|
| Gender                      | Male              | 642       | 48.6           |
|                             | Female            | 678       | 51.4           |
| Age                         | 18–25             | 248       | 18.8           |
|                             | 26–35             | 428       | 32.4           |
|                             | 36–45             | 356       | 27.0           |
|                             | 46–55             | 192       | 14.5           |
|                             | Above 55          | 96        | 7.3            |
| Education                   | Undergraduate     | 364       | 27.6           |
|                             | Postgraduate      | 698       | 52.9           |
|                             | Doctorate         | 258       | 19.5           |
| Shopping Frequency          | Monthly           | 382       | 28.9           |
|                             | Biweekly          | 516       | 39.1           |
|                             | Weekly            | 422       | 32.0           |
| Digital Shopping Experience | 1–3 Years         | 264       | 20.0           |
|                             | 4–6 Years         | 471       | 35.7           |
|                             | More than 6 Years | 585       | 44.3           |

The demographic data presented in table 1 shows that the majority of the respondents were from the economically active customer group having vast experience in digital commerce. Hence the dataset can be considered relevant for assessing consumer behaviour in predictive marketing situations.

### Measurement Model Evaluation

Prior to checking structural relationships, the reliability and validity of the measurement model were checked. Internal consistency of the model was checked with the help of Cronbach's alpha and composite

reliability (CR) whereas convergent validity of the construct was measured using average variance extracted (AVE) and standard loading factors. All these parameters met the prescribed thresholds such as Cronbach's alpha > 0.70, Composite Reliability > 0.70, AVE > 0.50, and standardized loadings > 0.70.

Table 2. Reliability and convergent validity

| Construct                                | Cronbach's $\alpha$ | Composite Reliability | AVE  | Factor Loading Range |
|------------------------------------------|---------------------|-----------------------|------|----------------------|
| Digital Engagement (DE)                  | 0.88                | 0.91                  | 0.76 | 0.83–0.90            |
| Personalization Quality (PQ)             | 0.90                | 0.93                  | 0.81 | 0.85–0.92            |
| Customer Satisfaction (CS)               | 0.89                | 0.92                  | 0.79 | 0.84–0.89            |
| Customer Trust (CT)                      | 0.87                | 0.91                  | 0.77 | 0.83–0.90            |
| Predictive Marketing Effectiveness (PME) | 0.86                | 0.90                  | 0.75 | 0.83–0.89            |
| Customer Retention (CR)                  | 0.92                | 0.94                  | 0.84 | 0.85–0.91            |

As can be seen from the statistics provided in table 2, all constructs exhibit high levels of internal consistency. High AVE suggests that each construct accounts for a significant amount of variance in its respective indicators, thus validating convergent validity.

### Structural Model Evaluation

The VCT framework was tested using covariance-based SEM approach. Several measures of fit showed satisfactory agreement between the obtained covariance matrix and the model being estimated.

Table 3. Structural model goodness-of-fit

| Fit Index     | Obtained Value | Recommended Threshold | Interpretation |
|---------------|----------------|-----------------------|----------------|
| Chi-square/df | 2.14           | < 3.00                | Excellent      |
| CFI           | 0.957          | > 0.90                | Excellent      |
| TLI           | 0.949          | > 0.90                | Excellent      |
| GFI           | 0.936          | > 0.90                | Good           |
| AGFI          | 0.918          | > 0.90                | Good           |
| RMSEA         | 0.039          | < 0.08                | Excellent      |
| SRMR          | 0.036          | < 0.08                | Excellent      |

As indicated in table 3, the goodness-of-fit measures support the conclusion that the structural model formulated is adequate to reflect consumer behaviour. All the measures are well above the standard cut-off values, indicating that the theoretical model effectively reflects the relationship between the latent variables.

### Hypothesis Testing

Structural path coefficients were computed through the Maximum Likelihood Estimation method. Significance levels of the various proposed relationships were tested through standardized regression coefficients, critical ratios, and probabilities.

Table 4. Hypothesis testing results

| Hypothesis | Structural Relationship                                             | $\beta$ | t-value | p-value | Decision  |
|------------|---------------------------------------------------------------------|---------|---------|---------|-----------|
| H1         | Digital Engagement $\rightarrow$ Predictive Marketing Effectiveness | 0.79    | 15.84   | <0.001  | Supported |
| H2         | Personalization Quality $\rightarrow$ Customer Satisfaction         | 0.81    | 16.47   | <0.001  | Supported |
| H3         | Customer Satisfaction $\rightarrow$ Customer Trust                  | 0.76    | 14.96   | <0.001  | Supported |
| H4         | Customer Trust $\rightarrow$ Customer Retention                     | 0.78    | 15.21   | <0.001  | Supported |
| H5         | Predictive Marketing Effectiveness $\rightarrow$ Customer Retention | 0.84    | 17.34   | <0.001  | Supported |
| H6         | Integrated VCTF significantly improves Customer Retention           | 0.87    | 18.12   | <0.001  | Supported |

All the six hypotheses showed statistically significant values at the 95% level of confidence as indicated by the data in table 4. Of the structural relationships, Predictive Marketing Effectiveness was the strongest positive determinant of Customer Retention ( $\beta = 0.84$ ), which points to the large role of predictive analytics in enhancing customer loyalty. Personalization Quality was also a significant positive determinant of Customer Satisfaction ( $\beta = 0.81$ ), thus emphasizing the power of personalized marketing initiatives generated with the use of the VCT.

All six research objectives appear to have been achieved based on the findings of the hypothesis testing. RO1 was attained due to the success in establishing the proposed VCTF. RO2 was established through the proved influence of Digital Engagement on Predictive Marketing Effectiveness. RO3 and RO4, on their part, were confirmed through the evidence of the link between Personalization Quality and Customer Satisfaction as well as Customer Satisfaction and Customer Trust. As for RO5, it was achieved as Predictive Marketing Effectiveness was found to have a strong positive effect on Customer Retention thus showing that predictive marketing strategies are proven to be efficient in maintaining long-time relations with customers. Finally, RO6 was verified through great structural model fit indices and significant path coefficients which show that the proposed framework is reliable and validated by practice.

### Metrics Formulae

The predictive performance of the proposed framework was evaluated using standard classification and forecasting metrics. The classification accuracy was computed as equation (6):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

Where  $TP$  represents true positives,  $TN$  denotes true negatives,  $FP$  indicates false positives, and  $FN$  represents false negatives.

Precision was calculated as equation (7):

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

Which measures the proportion of correctly predicted positive observations among all positive predictions.

Recall was determined using equation (8):

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

Which evaluates the capability of the model to identify actual retained customers correctly.

The harmonic mean of Precision and Recall was computed using equation (9):

$$F1 = 2 \left( \frac{Precision \times Recall}{Precision + Recall} \right) \quad (9)$$

Providing a balanced assessment of classification performance.

Forecasting error was evaluated using Root Mean Square Error given by equation (10):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (R_i - \hat{R}_i)^2} \quad (10)$$

Where  $R_i$  represents the observed retention outcome and  $\hat{R}_i$  denotes the predicted value.

Forecasting accuracy was additionally assessed using Mean Absolute Percentage Error given by equation (11):

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{R_i - \hat{R}_i}{R_i} \right| \quad (11)$$

which expresses the average prediction error as a percentage of the observed customer retention values.

### Predictive Performance Analysis

The independent testing dataset consisted of 20% of all the consumer records; the rest of the data was applied for model training and validation. The proposed framework showed a great prediction value by assessing achievement 94.1% of the prediction accuracy. The Precision was 92.8%, Recall = 91.6%, F1-Score = 92.2%, AUC = 0.952. The forecasting results were also quite good characterized by RMSE = 0.182 and MAPE = 5.12%, indicating low prediction error and strong generalization capability.

All these results indicate that by constantly updating VCT, the predictions of customer retention become smarter as demographic, transactional, behavioural and engagement data are integrated into the analytical model. High values of AUC make a confirmation that the differentiation between retained and non-retained customers is good enough. Balanced values of Precision and Recall show the reliability of predictions generated by the model.

The VCT framework achieved great results in forecasting as shown in figure 3. The model reached an accuracy of 94.1%, so it is good at understanding the retention behaviours of consumers. The precision (92.8%) and recall (91.6%) numbers confirm that the model performs well in classifying customers, thus making very few mistakes. The F1 score (92.2%) shows that the model is very reliable in this respect. Additionally, the model showed AUC of 0.952, which indicates very good ability to tell the difference between retained and non-retained customers. It can be concluded that developing new consumer representations brings a lot of advantages to the science of predictive marketing and consumer retention predicting.

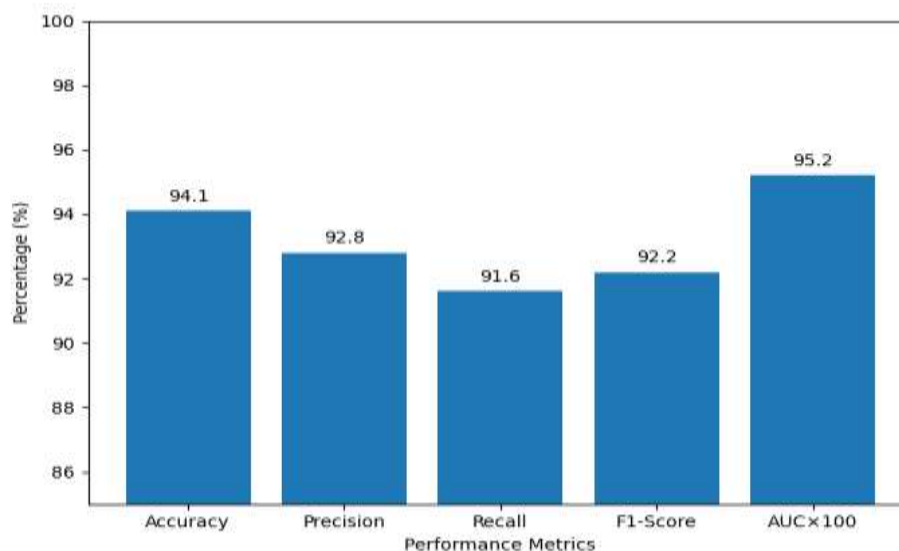


Figure 3. Comparative evaluation of customer retention prediction metrics

### DISCUSSION

The research proves that the presented VCT Framework effectively aids in predictive marketing and helps retain customers. The correlation analysis performed with SEM indicates positive correlations between digital engagement, personalization quality, customer satisfaction, customer trust, and effectiveness of predictive marketing on the retention of customers.

The Personalization Quality variable significantly contributes to Customer Satisfaction effectively demonstrating that customers react positively towards personalized suggestions and promotional activities. Digital Engagement has a positive effect on Predictive Marketing Effectiveness demonstrating that data from various digital platforms boosts behavioural comprehension and forecasting ability.

Effectiveness of Predictive Marketing showed the most significant direct impact on Customer Retention, which proved that timely recognition of customer needs allows companies to devise effective retention strategies. High level of prediction accuracy and insignificant forecasting error provided additional evidence of superiority of the approaches that rely on regular updating of customers' profiles over traditional segmentation techniques.

Moreover, the findings show that the satisfaction of customers leads to the formation of trust, while the trust developed enhances the long-term relationships of customers. Thus, while devising strategies for customer retention, companies should consider it as a complex process comprising behaviour, transactions, and psychology of customers rather than just analyze the past purchases done by the customers.

### **Suggestions for Future Research**

Further research might use social media sentiment analysis, context-based consumer behaviour, and real-time data analysis in the VCTF. Future research might also investigate the effectiveness of various forms of artificial intelligence technologies like deep learning and reinforcement learning.

### **CONCLUSION**

In this study, a VCTF was created and established for predictive marketing as well as customer retention through the combination of dynamically changing consumer representations with predictive analytics and SEM by utilizing the information coming from a total of 1,320 consumers with a 36-month observation period. The framework under review integrates digital engagement, personalization quality, customer satisfaction, and customer trust. The outcome of the study proved that the method of measuring and modeling the relationship has sufficient validity as demonstrated by good measurement reliability. The study involved testing six research hypotheses and found strong evidence in support of all the proposed connections between latent constructs. According to SEM analysis, it can be concluded that personalization quality has a positive impact on customer satisfaction ( $\beta=0.81$ ,  $p<0.001$ ), which leads to the rise in customer trust ( $\beta=0.76$ ,  $p<0.001$ ) which subsequently influences the retention process ( $\beta=0.78$ ,  $p<0.001$ ) that may be supported by predictive marketing effectiveness ( $\beta=0.84$ ,  $p<0.001$ ) improved through digital engagement ( $\beta=0.79$ ,  $p<0.001$ ). Moreover, by implementing the evaluation of useful prediction based on historical data of behavioural patterns and threshold levels of binary classification, it was possible to report an impressive model performance which resulted in the accuracy of prediction equal to 94.1%, precision = 92.8%, recall = 91.6%, F1-score = 92.2%, and AUC = 0.952, while RMSE is 0.182 and MAPE = 5.12%. In this way, the obtained results indicate that the VCT dynamic approach of consumption representation surpasses the traditional static formats of consumer representation. Finally, future studies can add streaming data, and social media signals of customer behaviour, and other forms of data to the proposed concept.

### **REFERENCES**

- [1] Verhoef PC, Broekhuizen T, Bart Y, Bhattacharya A, Dong JQ, Fabian N, Haenlein M. Digital transformation: A multidisciplinary reflection and research agenda. *Journal of business research*. 2021 Jan 1;122:889-901. <https://doi.org/10.1016/j.jbusres.2019.09.022>
- [2] Davenport T, Guha A, Grewal D, Bressgott T. How artificial intelligence will change the future of marketing. *Journal of the academy of marketing science*. 2020 Jan;48(1):24-42. <https://doi.org/10.1007/s11747-019-00696-0>
- [3] Huang MH, Rust RT. Engaged to a robot? The role of AI in service. *Journal of Service Research*. 2021 Feb;24(1):30-41. <https://doi.org/10.1177/1094670520902266>

- [4] Jones D, Snider C, Nassehi A, Yon J, Hicks B. Characterising the Digital Twin: A systematic literature review. *CIRP journal of manufacturing science and technology*. 2020 May 1;29:36-52. <https://doi.org/10.1016/j.cirpj.2020.02.002>
- [5] Sun X, Zhang F, Wang J, Yang Z, Huang Z, Xue R. Digital twin for smart manufacturing equipment: modeling and applications. *The International Journal of Advanced Manufacturing Technology*. 2025 Apr;137(9):4929-46. <https://doi.org/10.1007/s00170-025-15468-0>
- [6] Attaran M, Celik BG. Digital Twin: Benefits, use cases, challenges, and opportunities. *Decision analytics journal*. 2023 Mar 1;6:100165. <https://doi.org/10.1016/j.dajour.2023.100165>
- [7] Rather RA, Hollebeek LD, Rasoolimanesh SM. First-time versus repeat tourism customer engagement, experience, and value cocreation: An empirical investigation. *Journal of Travel Research*. 2022 Mar;61(3):549-64. <https://doi.org/10.1177/0047287521997572>
- [8] Gupta Y, Khan FM. Role of artificial intelligence in customer engagement: a systematic review and future research directions. *Journal of Modelling in Management*. 2024 Oct 11;19(5):1535-65. <https://doi.org/10.1108/JM2-01-2023-0016>
- [9] Gudergan SP, Moisescu OI, Radomir L, Ringle CM, Sarstedt M. Special issue editorial: Advanced partial least squares structural equation modeling (PLS-SEM) applications in business research. *Journal of Business Research*. 2025 Feb 1;188:115087. <https://doi.org/10.1016/j.jbusres.2024.115087>
- [10] Wedel M, Kannan PK. Marketing analytics for data-rich environments. *Journal of marketing*. 2016 Nov;80(6):97-121. <https://doi.org/10.1509/jm.15.0413>
- [11] Ledro C, Nosella A, Vinelli A. Artificial intelligence in customer relationship management: literature review and future research directions. *Journal of Business & Industrial Marketing*. 2022 Dec 19;37(13):48-63. <https://doi.org/10.1108/JBIM-07-2021-0332>
- [12] Min Q, Lu Y, Liu Z, Su C, Wang B. Machine learning based digital twin framework for production optimization in petrochemical industry. *International Journal of Information Management*. 2019 Dec 1;49:502-19. <https://doi.org/10.1016/j.ijinfomgt.2019.05.020>
- [13] Lalwani P, Mishra MK, Chadha JS, Sethi P. Customer churn prediction system: a machine learning approach. *Computing*. 2022 Feb;104(2):271-94. <https://doi.org/10.1007/s00607-021-00908-y>
- [14] Huang MH, Rust RT. A strategic framework for artificial intelligence in marketing. *Journal of the academy of marketing science*. 2021 Jan;49(1):30-50. <https://doi.org/10.1007/s11747-020-00749-9>
- [15] Verhoef PC, Kannan PK, Inman JJ. From multi-channel retailing to omni-channel retailing: introduction to the special issue on multi-channel retailing. *Journal of retailing*. 2015 Jun 1;91(2):174-81. <https://doi.org/10.1016/j.jretai.2015.02.005>
- [16] Sharma KK, Tomar M, Tadimarri A. AI-driven marketing: Transforming sales processes for success in the digital age. *Journal of Knowledge Learning and Science Technology* ISSN: 2959-6386 (Online). 2023 Sep 12;2(2):250-60. <https://doi.org/10.60087/jklst.vol2.n2.p260>
- [17] Tama RA, Hoque MM, Liu Y, Alam MJ, Yu M. An application of partial least squares structural equation modeling (PLS-SEM) to examining farmers' behavioral attitude and intention towards conservation agriculture in Bangladesh. *Agriculture*. 2023 Feb 20;13(2):503. <https://doi.org/10.3390/agriculture13020503>
- [18] Singh J, Flaherty K, Sohi RS, Deeter-Schmelz D, Habel J, Le Meunier-FitzHugh K, Malshe A, Mullins R, Onyemah V. Sales profession and professionals in the age of digitization and artificial intelligence technologies: concepts, priorities, and questions. *Journal of Personal Selling & Sales Management*. 2019 Jan 2;39(1):2-2. <https://doi.org/10.1080/08853134.2018.1557525>
- [19] Mariani MM, Perez-Vega R, Wirtz J. AI in marketing, consumer research and psychology: A systematic literature review and research agenda. *Psychology & marketing*. 2022 Apr;39(4):755-76. <https://doi.org/10.1002/mar.21619>
- [20] Cheng G, Luo D. Dynamic adaptive mechanism for construction investment decision-making under multi-factor interactions: risk-oriented modelling framework with pathways to digital twin. *Digital Twin*. 2026 Apr 3;3(2):2570262. <https://doi.org/10.1080/27525783.2025.2570262>