WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Generative AI for Automating Consumer Behaviour Prediction in Marketing Strategies

Mrs. Gorisetty Nirosha [1], Pothula Rama Sri [2]

Department of Computer Science and Engineering
GMR Institute of Technology, Rajam, Andhra Pradesh, India

ABSTRACT—

Consumer behaviour plays an important role in developing effective marketing strategies. Understanding customer preferences, purchasing patterns, product interests and feedback helps businesses make better marketing decisions. Traditional marketing approaches often depend on manual analysis and basic statistical techniques, which may not effectively handle large volumes of customer data. Recent developments in Artificial Intelligence (AI), Machine Learning (ML) and Generative AI provide opportunities to automate consumer behaviour analysis and support personalized marketing.

This paper proposes an integrated framework for consumer behaviour prediction using Machine Learning and Generative AI techniques. The proposed framework combines purchase behaviour prediction, customer segmentation, sentiment analysis and automated marketing insight generation. XGBoost is considered for predicting purchase behaviour, while K-Means and Hierarchical Clustering are considered for customer segmentation. Sentiment analysis is used to understand customer opinions from reviews and feedback. Generative AI is then used to convert the analytical outputs into meaningful customer insights and personalized marketing suggestions. The proposed approach aims to connect predictive analytics with automated marketing decision support. The framework is discussed in terms of its architecture, methodology, potential applications, expected outcomes and limitations. It provides a foundation for developing more data-driven and personalized marketing strategies.

Keywords: Generative AI, Consumer Behaviour Prediction, Machine Learning, XGBoost, Customer Segmentation, Sentiment Analysis, Personalized Marketing, Predictive Analytics.

1. INTRODUCTION

Marketing strategies have changed significantly with the growth of digital platforms and online shopping. Customers interact with businesses through websites, mobile applications, social media platforms and e-commerce services. These interactions generate large amounts of data, including browsing history, product views, purchase records, customer reviews and feedback. Understanding this information helps organizations identify customer preferences and predict future purchasing behaviour. However, analyzing large and continuously growing customer datasets manually is difficult. Traditional marketing methods may also fail to identify complex relationships between customer interests and purchasing decisions. Artificial Intelligence and Machine Learning provide techniques for analyzing customer data and identifying useful patterns. Machine Learning algorithms can predict customer purchase intentions, classify customer groups and identify factors influencing purchasing decisions. Sentiment analysis can further help organizations understand customer opinions by identifying positive, negative and neutral feedback. Generative AI introduces another opportunity by generating meaningful information from analytical results. Instead of presenting only numerical predictions and charts, Generative AI can convert the outputs of Machine Learning models into understandable insights and personalized marketing suggestions. This paper proposes an integrated framework that combines Machine Learning, customer segmentation, sentiment analysis and Generative AI to support automated consumer behaviour prediction and marketing strategy development.

1.1 Background

Consumer behaviour refers to the actions and decisions of individuals when selecting, purchasing and using products or services. It is influenced by several factors, including personal preferences, product quality, price, customer experience, advertisements and social influence.

With the availability of digital customer data, organizations can apply predictive analytics to understand these factors. Machine Learning techniques such as Decision Trees, Random Forest, Support Vector Machines and XGBoost can be used to identify purchasing patterns and predict customer decisions.

Customer segmentation techniques group customers based on similarities in their behaviour or characteristics. Sentiment analysis identifies opinions from textual feedback. These analytical methods provide useful information for marketing activities.

Generative AI can use these outputs to produce understandable summaries, recommendations and personalized marketing content. Integrating these methods can support a more automated and data-driven marketing process.

1.2 Problem Statement

Businesses collect customer information from different digital platforms, but extracting meaningful marketing insights from this information remains challenging. Individual analytical techniques generally address specific tasks, such as purchase prediction, sentiment classification or customer segmentation.

When these techniques operate separately, organizations may need additional manual effort to combine their outputs and convert them into marketing decisions.

Therefore, an integrated framework is required to analyze consumer behaviour, identify customer groups, understand customer opinions and generate personalized marketing insights using Machine Learning and Generative AI.

1.3 Objectives

The main objectives of this paper are:

1. To study the application of Machine Learning techniques in consumer behaviour prediction.
2. To predict customer purchasing behaviour using classification algorithms such as XGBoost.
3. To identify customer groups using clustering techniques.
4. To analyze customer opinions and feedback using sentiment analysis.
5. To integrate Generative AI for generating personalized marketing insights.
6. To propose an automated framework that connects predictive analytics with marketing strategy development.

1.4 Scope of the Study

The proposed framework focuses on consumer behaviour analysis using customer-related data such as browsing activities, product interactions, purchasing information and customer feedback.

It considers purchase prediction, customer segmentation and sentiment analysis as the main analytical components. Generative AI is proposed as an additional component for converting analytical outputs into understandable marketing suggestions.

The study focuses on the design and proposed methodology of the framework. Its performance must be validated through implementation and experimental evaluation before making claims about actual predictive accuracy or business impact.

2. LITERATURE REVIEW

Several research studies have explored the use of Artificial Intelligence, Machine Learning and predictive analytics in understanding customer behaviour. The reviewed literature covers consumer sentiment, purchasing behaviour, customer segmentation and AI-based marketing personalization.

2.1 Consumer Behaviour Analysis Using Neurobehavioural and Social Media Sentiment Analysis

Najafabadi et al. (2024) investigated consumer behaviour using neurobehavioural information and social media sentiment analysis. The study considered customer responses to different products and examined how behavioural and sentiment-related information could support product development decisions.

The research demonstrates the importance of combining different sources of consumer information. It also highlights how customer opinions and behavioural responses can provide useful information for understanding preferences.

2.2 AI-Driven Marketing Personalization

The reviewed literature on AI-driven marketing personalization discusses the relationship between organizational capabilities and consumer acceptance. It considers factors such as organizational readiness, marketing competencies, consumer trust and privacy concerns.

The study highlights that implementing AI-based personalization requires more than technical capabilities. Organizations must also consider customer acceptance, responsible data usage and transparency when developing personalized marketing strategies.

2.3 Consumer Online Shopping Behaviour Prediction Using Machine Learning

Chen (2025) explored consumer online shopping behaviour prediction using Machine Learning algorithms. The study focuses on analyzing online shopping-related information to identify patterns associated with purchasing behaviour.

XGBoost is an important algorithm considered in this area because it combines multiple decision trees through boosting to improve classification performance. The reviewed report mentions an accuracy value of approximately 0.91 for the referenced study. This value belongs to the referenced research and is not an experimental result of the proposed framework.

2.4 Customer Segmentation Using Data Analytics

The reviewed literature also discusses clustering-based segmentation using smart-meter and behavioural datasets. Techniques such as Hierarchical

Clustering, Self-Organizing Maps and K-Means are used to identify groups with similar characteristics.

Although smart-meter datasets are not directly equivalent to e-commerce customer datasets, the clustering methods demonstrate how unsupervised learning can identify meaningful groups from behavioural information.

2.5 Predictive Analytics for Customer Behaviour

GhorbanTanhaei et al. (2024) discussed predictive analytics for anticipating customer trends and preferences. The study considers Machine Learning methods such as Decision Trees, Random Forest, Logistic Regression, Support Vector Machines and Gradient Boosting.

The literature emphasizes the importance of evaluating predictive models using multiple measures, including Precision, Recall, F1-score and ROC-AUC, rather than depending only on accuracy.

2.6 Summary of Literature Review

The reviewed studies demonstrate that Machine Learning, sentiment analysis and clustering techniques can contribute to consumer behaviour analysis. However, the reviewed approaches focus on different analytical tasks. Integrating their outputs with Generative AI provides an opportunity to develop a more connected approach to marketing decision support.

3. RESEARCH GAP

The literature review identifies several areas that can be further explored.

1. Separate analytical processes: Purchase prediction, sentiment analysis and customer segmentation are often studied as individual tasks.
2. Limited integration: Combining different analytical outputs into a single marketing-support framework remains an important area for development.
3. Manual interpretation: Predictive results may require additional human interpretation before they can be converted into marketing actions.
4. Personalization challenges: Customer preferences and feedback may change over time, requiring marketing strategies to adapt to new information.
5. Responsible AI requirements: Privacy, transparency and customer trust must be considered when using personal information for automated marketing.

To address these areas, this paper proposes a framework that connects Machine Learning-based consumer analysis with Generative AI-based marketing insight generation.

4. PROPOSED METHODOLOGY

The proposed methodology combines data preprocessing, purchase behaviour prediction, customer segmentation, sentiment analysis and Generative AI.

4.1 Data Collection

The first stage involves collecting consumer-related information from available datasets or digital platforms. The information may include:

- Customer browsing history.
- Product views and product interaction records.
- Product-related session duration.
- Purchase history.
- Customer reviews and feedback.
- Relevant customer and product attributes.

The collected information forms the input for the subsequent analytical stages. Data collection should follow applicable privacy and data-protection requirements.

4.2 Data Preprocessing

Raw consumer data may contain missing values, duplicate records, inconsistent formats and irrelevant attributes. Preprocessing is therefore required before applying Machine Learning algorithms.

The preprocessing stage includes:

1. Handling missing values.
2. Removing duplicate records.
3. Identifying and treating outliers where appropriate.
4. Encoding categorical variables.
5. Scaling numerical features when required.
6. Selecting relevant features for model training.

Feature engineering is also used to create useful variables from existing data. For example, the number of product views and the duration of product-

related activities may help represent customer interest.

4.3 Consumer Purchasing Behaviour Prediction

The purchase prediction component aims to identify whether a customer is likely to complete a purchase based on available behavioural information.

XGBoost is considered as the primary classification algorithm. It is an ensemble learning method that builds decision trees sequentially, where each new tree attempts to improve the errors made by previous trees.

The model receives preprocessed customer-related features and produces a purchase prediction.

The general workflow is:

1. Select relevant consumer behaviour features.
2. Divide the dataset into training and testing subsets.
3. Train the XGBoost classification model.
4. Predict purchasing behaviour on the testing subset.
5. Evaluate the model using suitable performance metrics.

Other classification algorithms, such as Decision Tree, Random Forest, Support Vector Machine and Gradient Boosting, may also be considered for comparison.

4.4 Customer Segmentation

Customer segmentation is used to group customers with similar behavioural characteristics. K-Means and Hierarchical Clustering are considered for this stage.

K-Means Clustering: K-Means divides the dataset into a predefined number of clusters. It assigns each customer to the cluster with the nearest centroid and repeatedly updates the centroids until the clustering process converges.

Hierarchical Clustering: Hierarchical Clustering creates a hierarchy of customer groups based on similarities between observations. The resulting hierarchy can be represented using a dendrogram.

The resulting customer groups may represent different behavioural patterns, such as frequent product viewers, occasional shoppers and customers with stronger purchasing activity. These descriptions are illustrative; actual segments depend on the dataset and clustering results.

4.5 Sentiment Analysis

Customer reviews and feedback contain information about product satisfaction, service quality and customer expectations.

Sentiment analysis is used to classify textual feedback into categories such as:

- Positive sentiment.
- Negative sentiment.
- Neutral sentiment.

The resulting sentiment information can help identify common customer opinions and product-related concerns. These outputs can be combined with purchasing behaviour and customer segmentation information to provide a broader understanding of consumer preferences.

4.6 Generative AI-Based Marketing Insight Generation

The Generative AI component uses the outputs of the analytical modules to generate understandable marketing insights.

Its inputs may include:

- Purchase prediction results.
- Customer segment information.
- Sentiment analysis results.
- Relevant product and customer attributes.

Based on these inputs, the Generative AI module can generate suggestions such as personalized product recommendations, customer engagement messages and marketing strategy ideas.

For example, if a customer segment shows interest in a particular product category and the feedback indicates positive sentiment, the system may generate a suggestion for a targeted promotional campaign related to that category.

Generative AI acts as an insight-generation component. Its suggestions should be reviewed before important marketing decisions are implemented.

5. SYSTEM DESIGN AND ARCHITECTURE

The proposed system consists of five major components: the web application, application backend, database, Machine Learning prediction module and Generative AI module.

5.1 System Architecture

The architecture follows a modular design in which each component performs a specific responsibility.

1. Web Application (Frontend):

Provides the interface for product browsing, customer interaction and feedback submission.

2. Application Backend:

Handles application requests, business logic, data processing and communication between the frontend, database and analytical modules. Flask with Python is considered for backend development.

3. Database:

Stores relevant information such as consumer profiles, product details, browsing activities and customer feedback.

4. Machine Learning Prediction Module:

Processes the required features and performs purchase behaviour prediction, customer segmentation and sentiment analysis.

5. Generative AI Module:

Uses the analytical outputs to generate marketing insights, personalized recommendations and customer engagement suggestions.

5.2 Working Process

The proposed system follows these steps:

1. Customers interact with the web application.
2. Relevant browsing, purchasing and feedback information is collected.
3. The backend stores and prepares the required data.
4. The Machine Learning module analyzes consumer behaviour.
5. Customer groups and sentiment categories are identified.
6. Generative AI converts the analytical outputs into marketing insights.
7. The generated suggestions are presented to support marketing decisions.

5.3 Algorithm Selection

The following algorithms and techniques are considered in the proposed framework.

Technique	Purpose
XGBoost	Purchase behaviour prediction
Random Forest	Comparative classification
Decision Tree	Classification and interpretable decisions
Support Vector Machine	Classification
K-Means	Customer segmentation
Hierarchical Clustering	Hierarchical customer grouping
Sentiment Analysis	Customer feedback classification
Generative AI	Marketing insight generation

The techniques are selected according to their intended roles. Their actual suitability and performance must be established through experimental evaluation.

6. IMPLEMENTATION APPROACH

The proposed framework can be implemented using Python and commonly available Machine Learning libraries.

6.1 Tools and Technologies

Component	Technology
Programming language	Python 3.x
Development environment	Visual Studio Code / Jupyter Notebook
Data processing	Pandas, NumPy
Machine Learning	Scikit-learn, XGBoost
Data visualization	Matplotlib, Seaborn
Backend	Flask
Database	Relational database
Generative AI	Suitable Generative AI model or API

6.2 Model Development Process

The implementation process can be organized into the following stages:

1. Import the required Python libraries.
2. Load the consumer behaviour dataset.
3. Perform data cleaning and preprocessing.
4. Select relevant input features and target variables.
5. Divide the dataset into training and testing subsets.
6. Train the selected classification algorithms.
7. Generate purchase behaviour predictions.
8. Apply clustering techniques to identify customer segments.
9. Perform sentiment analysis on customer feedback.
10. Provide the analytical outputs to the Generative AI component.
11. Generate and review marketing insights.

6.3 Illustrative XGBoost Implementation

The following code illustrates the purchase prediction process using XGBoost. It is a sample implementation and does not represent a verified experimental result.

```
import pandas as pd

from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
from xgboost import XGBClassifier

# Load the consumer dataset
data = pd.read_csv("consumer_data.csv")

# Select input features
X = data[
    [
        "PageValues",
        "ProductRelated",
        "ProductRelated_Duration"
    ]
]

# Select target variable
y = data["Revenue"]

# Split the dataset
X_train, X_test, y_train, y_test = train_test_split(
    X,
    y,
    test_size=0.2,
    random_state=42
)

# Train the XGBoost model
model = XGBClassifier()
model.fit(X_train, y_train)

# Generate predictions
prediction = model.predict(X_test)

# Evaluate model
accuracy = accuracy_score(y_test, prediction)

print("Accuracy:", accuracy)
```

The dataset must contain the specified columns and suitable target labels for this example to run. Additional preprocessing and class-imbalance handling may be required depending on the actual dataset.

Figure 1. Proposed System Architecture

WEB APPLICATION (FRONTEND) **Product browsing | Customer interaction | Feedback**
APPLICATION BACKEND (FLASK / PYTHON) **API handling | Business logic | Data preparation**
DATABASE **Consumer profiles | Products | Activities | Feedback**
MACHINE LEARNING MODULE **Purchase prediction | Customer segmentation | Sentiment analysis**
GENERATIVE AI MODULE **Marketing insights | Personalized suggestions | Engagement content**

7. RESULTS AND DISCUSSION

7.1 Expected Outcomes

The proposed framework is expected to support the following outcomes:

1. Purchase Behaviour Prediction: The classification model is expected to identify purchasing patterns from customer interaction data.
2. Customer Segmentation: Clustering techniques are expected to group customers according to similarities in their behavioural features.
3. Sentiment Identification: Sentiment analysis is expected to classify customer feedback into positive, negative and neutral categories.
4. Marketing Insight Generation: Generative AI is expected to convert analytical outputs into understandable marketing suggestions.
5. Integrated Decision Support: Combining the analytical components is expected to reduce the need for separate manual interpretation of different outputs.

These are expected outcomes of the proposed framework, not experimentally established findings.

7.2 Evaluation Metrics

The performance of the purchase prediction model can be evaluated using the following metrics:

Accuracy: Measures the proportion of correctly classified observations.

Accuracy=\frac{TP+TN}{TP+TN+FP+FN}

Precision: Measures the proportion of predicted positive cases that are actually positive.

Precision=\frac{TP}{TP+FP}

Recall: Measures the proportion of actual positive cases correctly identified by the model.

Recall=\frac{TP}{TP+FN}

F1-Score: Represents the harmonic mean of Precision and Recall.

F1=2\times\frac{Precision\times Recall}{Precision+Recall}

Here, TP represents True Positives, TN represents True Negatives, FP represents False Positives and FN represents False Negatives.

For customer segmentation, suitable measures such as the Silhouette Score can be considered. Sentiment classification can be evaluated using Accuracy, Precision, Recall and F1-score when labelled test data is available.

7.3 Discussion

The proposed framework combines multiple analytical techniques to understand consumer behaviour from different perspectives. Purchase prediction provides information about possible purchasing decisions, while customer segmentation identifies groups with similar characteristics. Sentiment analysis adds information about customer opinions and experiences.

Generative AI provides a way to convert these outputs into readable marketing suggestions. This integration can support personalized marketing activities and help organizations interpret customer-related information more efficiently.

However, the quality of the generated insights depends on the quality of the input data, the performance of the analytical models and the reliability of the Generative AI component. Therefore, model evaluation, human review and responsible data handling are important before using the framework in real-world marketing environments.

8. ADVANTAGES AND LIMITATIONS

8.1 Advantages

- Combines purchase prediction, customer segmentation and sentiment analysis.
- Supports data-driven marketing decisions.
- Provides a way to generate personalized marketing suggestions.
- Connects Machine Learning outputs with understandable insights.
- Uses modular components that can be developed and evaluated separately.

8.2 Limitations

- Prediction quality depends on dataset quality and representativeness.
- Imbalanced customer data may affect classification performance.
- Customer preferences may change over time.
- Generative AI may produce inaccurate or unsuitable suggestions.
- Privacy and responsible use of customer information must be addressed.
- Real-world effectiveness requires implementation and experimental validation.

9. FUTURE SCOPE

The proposed framework can be extended in several directions.

1. Real-Time Prediction: Integrating real-time customer interactions from e-commerce websites and digital platforms.
2. Advanced Generative AI Models: Using large language models to generate more context-aware marketing insights.
3. Adaptive Learning: Updating prediction models as new customer behaviour data becomes available.
4. Multimodal Data Analysis: Combining textual reviews, images and other relevant customer interaction data.
5. Explainable AI: Providing explanations for predictions and recommendations to improve transparency.
6. Privacy-Aware Marketing: Applying suitable privacy-preserving methods and responsible data management practices.

Future research can also evaluate the proposed framework using real-world datasets and compare different Machine Learning algorithms under consistent experimental conditions.

10. CONCLUSION

This paper proposed an integrated framework for automating consumer behaviour prediction in marketing strategies using Machine Learning and Generative AI. The framework combines purchase behaviour prediction, customer segmentation and sentiment analysis to identify useful information from consumer-related data. XGBoost is considered for purchase prediction, while K-Means and Hierarchical Clustering are considered for customer segmentation. Sentiment analysis helps identify customer opinions from reviews and feedback. Generative AI is proposed to convert the outputs of these analytical components into understandable insights and personalized marketing suggestions.

The main contribution of the proposed approach is the integration of different consumer analysis techniques with Generative AI-based insight generation. Rather than treating purchase prediction, segmentation and sentiment analysis as completely separate tasks, the framework connects their outputs to support marketing decision-making. The proposed framework provides a foundation for developing automated and personalized marketing systems. Further implementation and experimental evaluation are required to measure its actual predictive performance and practical effectiveness.

ACKNOWLEDGEMENT

The author expresses sincere gratitude to the faculty members of the Department of Computer Science and Engineering, GMR Institute of Technology, for their guidance and support during the preparation of this work.

REFERENCES

[1] Najafabadi, A. J., Skryzhadlovska, A., & Valilai, O. F. (2024). Agile product development by prediction of consumers' behaviour; using neurobehavioral and social media sentiment analysis approaches. *Procedia Computer Science*, *232*, 1683-1693.

[2] Chen, X. (2025). Consumer Online Shopping Behavior Prediction Based on Machine Learning Algorithm. *Procedia Computer Science*, *262*, 1395-1401.

[3] Lee, S., Pu, Z., Goetzinger, G., Preisach, C., Haft, M., Southall, A., & Vogel, A. (2025). *U.S. Patent Application No. 13/239,922*.

[4] GhorbanTanhaei, H., Boozary, P., Sheykhan, S., Rabiee, M., Rahmani, F., & Hosseini, I. (2024). Predictive analytics in customer behavior: Anticipating trends and preferences. *Results in Control and Optimization*, *17*, 100462.

[5] Cai, B. (2024). Research on consumer behavior and prediction based on deep learning in the new energy vehicle market. *Procedia Computer Science*, *247*, 351-357.

[6] Lajoie, P. L. (2025). Content marketing as a propaganda vehicle for a romantic-managerial conception of artificial intelligence. *Critical Perspectives on Accounting*, *102*, 102810.

[7] Lafuente, S. G., Prados-Castillo, J. F., Guaita-Martínez, J. M., & Ribeiro-Soriano, D. (2026). Digital transformation, artificial intelligence, and sustainable marketing strategies in cooperatives. *Sustainable Technology and Entrepreneurship*, 100138.

[8] Dong, X., Feng, Z., Yu, X., Zhang, J., & Jia, H. (2024). Evolution prediction method for electric private car ownership considering the decision-making behavior of consumers. *Energy Reports*, *11*, 498-509.

[9] Prakash, K. P., Kumar, Y. P., Reddy, C. P., Pradeep, D. J., Flah, A., Alzaed, A. N., ... & Ghoneim, S. S. (2025). A comprehensive analytical exploration and customer behavior analysis of smart home energy consumption data with a practical case study. *Energy Reports*, *8*, 9081-9093.

[10] Wuyart, A., & Millot, J. M. (2026). Generative AI for CSF multi-omics: Learning strategies for small-cohort classification. *Neuroscience Informatics*, 100278.