Data and Metadata. 2025; 4:700
doi: 10.56294/dm2025700

ORIGINAL

Decoding Consumer Behaviour: Leveraging Big Data and Machine Learning for Personalized Digital Marketing

Decodificando el comportamiento del consumidor: aprovechando Big Data y el aprendizaje automático para un marketing digital personalizado

Anber Abraheem Shlash Mohammad[1], Suleiman Ibrahim Shelash Mohammad[2,3], Badrea Al Oraini[4], Ayman Hindieh[5], Asokan Vasudevan[6], Muhammad Turki Alshurideh[7]

[1]Digital Marketing Department, Faculty of Administrative and Financial Sciences, Petra University, Jordan.
[2]Electronic Marketing and Social Media, Economic and Administrative Sciences Zarqa University, Jordan.
[3]Research follower, INTI International University, 71800 Negeri Sembilan, Malaysia.
[4]Department of Business Administration, Collage of Business and Economics, Qassim University, Qassim - Saudi Arabia.
[5]Electronic Marketing and Social Media, Economic and Administrative Sciences Zarqa University, Jordan.
[6]Faculty of Business and Communications, INTI International University, 71800 Negeri Sembilan, Malaysia.
[7]Department of Marketing, School of Business, The University of Jordan, Amman 11942, Jordan.

Cite as: Shlash Mohammad AA, Shelash Mohammad SI, Al Oraini B, Ayman Hindieh AH, Asokan Vasudevan AV, Turki Alshurideh M. Decoding Consumer Behaviour: Leveraging Big Data and Machine Learning for Personalized Digital Marketing. Data and Metadata.2025; 4:700. https://doi.org/10.56294/dm2025700

Submitted: 26-02-2024 **Revised:** 01-08-2024 **Accepted:** 02-01-2025 **Published:** 03-01-2025

Editor: Dr. Adrián Alejandro Vitón Castillo

Corresponding author: Anber Abraheem Shlash Mohammad

ABSTRACT

Introduction: big data analytics and machine learning have transformed digital marketing by enabling data-driven insights for personalization. This study investigates the role of engagement metrics, sentiment analysis, and consumer segmentation in enhancing marketing effectiveness. Specifically, it examines how these technologies process consumer interaction data to uncover actionable insights, segment audiences, and drive purchase conversions.
Method: the study employed a mixed-methods approach, integrating big data analytics and machine learning techniques. Descriptive statistics highlighted engagement patterns, while k-means clustering segmented consumers based on behavioural and emotional data. Sentiment analysis, conducted using Natural Language Processing (NLP), captured consumer emotions as positive, neutral, or negative. Regression analysis evaluated the influence of social media activity, click-through rates, session duration, and sentiment scores on purchase conversion rates.
Results: descriptive analysis revealed significant variability in consumer engagement and sentiment, with 37,5 % of consumers expressing positive sentiment. Clustering identified three distinct consumer segments, reflecting differences in engagement and sentiment. Regression analysis showed that sentiment had a positive but statistically insignificant relationship with purchase conversions, while other metrics, such as click-through rates and session duration, exhibited minimal impact. The overall explanatory power of the regression model was low (R-squared = 0,001), indicating the need for additional factors to understand purchase behaviour.
Conclusion: the findings emphasize the potential of big data analytics and machine learning in consumer segmentation and sentiment analysis. However, their direct impact on purchase conversion is limited without integrating broader variables. A holistic, adaptive framework combining behavioural, emotional, and contextual insights is essential for maximizing marketing personalization and driving outcomes in dynamic digital environments.

Keywords: Consumer Behaviour; Data and Metadata; Machine Learning; Digital Marketing; Personalisation.

RESUMEN

Introducción: el análisis de big data y el aprendizaje automático han transformado el marketing digital al permitir la obtención de información basada en datos para la personalización. Este estudio investiga el papel de las métricas de interacción, el análisis de sentimientos y la segmentación de consumidores en la mejora de la eficacia del marketing. En concreto, examina cómo estas tecnologías procesan los datos de interacción de los consumidores para descubrir información útil, segmentar audiencias e impulsar las conversiones de compra.
Método: el estudio empleó un enfoque de métodos mixtos, integrando técnicas de análisis de big data y aprendizaje automático. Las estadísticas descriptivas destacaron los patrones de interacción, mientras que la agrupación de k-medias segmentó a los consumidores en función de los datos emocionales y de comportamiento. El análisis de sentimientos, realizado mediante el procesamiento del lenguaje natural (PLN), capturó las emociones de los consumidores como positivas, neutrales o negativas. El análisis de regresión evaluó la influencia de la actividad en las redes sociales, las tasas de clics, la duración de la sesión y las puntuaciones de sentimiento en las tasas de conversión de compra.
Resultados: el análisis descriptivo reveló una variabilidad significativa en la participación y el sentimiento de los consumidores, con un 37,5 % de los consumidores expresando un sentimiento positivo. La agrupación identificó tres segmentos de consumidores distintos, lo que refleja diferencias en la participación y el sentimiento. El análisis de regresión mostró que el sentimiento tenía una relación positiva pero estadísticamente insignificante con las conversiones de compra, mientras que otras métricas, como las tasas de clics y la duración de la sesión, exhibieron un impacto mínimo. El poder explicativo general del modelo de regresión fue bajo (R cuadrado = 0,001), lo que indica la necesidad de factores adicionales para comprender el comportamiento de compra.
Conclusiones: los hallazgos enfatizan el potencial del análisis de big data y el aprendizaje automático en la segmentación de consumidores y el análisis de sentimientos. Sin embargo, su impacto directo en la conversión de compra es limitado sin la integración de variables más amplias. Un marco holístico y adaptativo que combine información conductual, emocional y contextual es esencial para maximizar la personalización del marketing e impulsar resultados en entornos digitales dinámicos.

Palabras clave: Comportamiento del Consumidor; Datos y Metadatos; Aprendizaje Automático; Marketing Digital; Personalización.

INTRODUCTION

The remarkable proliferation of digital technologies has fundamentally reshaped the marketing landscape, transforming the way businesses engage with consumers.[1,2,3] At the heart of this evolution lies the integration of big data analytics and machine learning, which have emerged as critical tools for deciphering consumer behaviour and personalizing marketing strategies.[4,5,6,7] These technologies allow businesses to process vast datasets, uncover actionable insights, and deliver tailored marketing campaigns that resonate with individual consumers. This research investigated the intersection of big data analytics, machine learning, and consumer behaviour, focused on their combined potential to enhance marketing personalization and drive better outcomes in a competitive digital environment.

Big data analytics has fundamentally changed the way organizations process consumer interaction data, which originates from various touchpoints such as social media platforms, e-commerce websites, and clickstream logs.[1,8,9] This interaction data is often complex, consisting of structured, semi-structured, and unstructured formats. The ability to process and analyse such data provides businesses with insights into consumer preferences, behaviours, and trends. When paired with machine learning, these insights become even more potent, as machine learning algorithms can identify hidden patterns, predict future trends, and segment consumers based on behavioural and emotional attributes.[10] This combination allows marketers to shift from generic, one-size-fits-all campaigns to dynamic, personalized strategies that adapt to individual consumer needs in real time.

The importance of this research field stems from the increasing consumer demand for personalized experiences. Modern consumers expect brands to understand their preferences and deliver relevant content and offers. Failure to meet these expectations can lead to disengagement and loss of customer loyalty. Personalization in marketing has been shown to significantly enhance consumer satisfaction, engagement, and conversion rates.[11] However, achieving effective personalization requires businesses to overcome several challenges, including data silos, the complexity of consumer behaviour, and ethical considerations related to data privacy.[11,12,13] This study addresses these challenges by exploring how big data analytics and machine learning can process consumer interaction data to generate actionable insights for marketing personalization.

The geographic focus of this study is Jordan, a country experiencing rapid digital transformation and increasing adoption of e-commerce and social media platforms. Jordan's youthful population, with over 60 % under the age of 30, is highly active online, presenting a significant opportunity for businesses to leverage digital marketing strategies tailored to local consumer behavior.[14] The country's digital ecosystem has seen substantial growth, with expanding internet penetration, rising smartphone usage, and a growing reliance on social media as a primary channel for brand-consumer interactions. However, despite these advancements, the application of big data analytics and machine learning in marketing remains relatively underexplored in the Jordanian context.[15] By examining the effectiveness of these technologies in processing consumer interaction data and enhancing personalization, this study not only addresses local challenges but also provides insights that can inform marketing strategies in emerging markets with similar digital landscapes. This focus on Jordan underscores the need for region-specific research to bridge the gap between global advancements in digital marketing and their local applications.

Despite its transformative potential, the integration of big data analytics and machine learning in marketing is still fraught with challenges. One of the primary gaps in the literature is the limited understanding of how these technologies influence key marketing outcomes such as engagement, sentiment, and purchase conversions.[16,17,18] While previous studies have demonstrated the utility of engagement metrics and sentiment analysis, few have examined their combined impact when processed through machine learning-driven consumer segmentation models.[19,20,21] Additionally, the direct relationship between consumer behaviour analytics and marketing personalization remains underexplored, particularly in dynamic and real-time digital environments. This research seeks to fill these gaps by investigating how advanced analytics and machine learning techniques can enhance personalization strategies in digital marketing.

The significance of this study lies in its potential to provide both theoretical insights and practical applications. From a theoretical perspective, the study contributes to the growing body of knowledge on consumer behaviour analytics by integrating traditional engagement metrics with advanced machine learning techniques. It explores the role of sentiment analysis in capturing consumer emotions and examines how these emotions influence marketing outcomes.[22] From a practical standpoint, the research offers actionable recommendations for businesses seeking to leverage big data and machine learning to optimize their marketing strategies. By providing a scalable framework for processing consumer interaction data, the study addresses the pressing need for effective tools to navigate the complexities of digital marketing.

The novelty of this study lies in its holistic approach to analysing consumer behaviour. Unlike previous research that focuses on isolated metrics or techniques, this study adopts a comprehensive framework that combines engagement metrics, sentiment analysis, and machine learning-driven segmentation. By integrating these components, the research provides a deeper understanding of the interplay between consumer behaviour and marketing outcomes. It highlights the importance of real-time adaptability in digital marketing strategies, emphasizing the need for dynamic, data-driven approaches to meet evolving consumer preferences.[1]

A central theme of this research is the transformative potential of sentiment analysis. Sentiment analysis, powered by natural language processing (NLP), allows businesses to capture consumer emotions as expressed in online interactions, such as social media comments and product reviews.[23] These emotions provide critical context for understanding consumer behaviour, as they often influence decision-making processes. Positive sentiment, for example, has been linked to higher levels of trust and purchase intent, while negative sentiment signals dissatisfaction that can hinder conversions.[24] This study examines the predictive power of sentiment analysis in marketing personalization, exploring how it complements engagement metrics and segmentation techniques to improve marketing outcomes.

Another key aspect of the research is the application of machine learning algorithms, such as k-means clustering, for consumer segmentation. Traditional segmentation methods often rely on demographic or psychographic data, which may not fully capture the complexity of consumer behaviour in digital environments. Machine learning offers a more sophisticated approach by analysing behavioural and emotional data to identify distinct consumer segments. These segments provide a granular understanding of consumer preferences and needs, enabling businesses to design targeted campaigns that resonate with specific audiences.[24] By integrating machine learning-driven segmentation with sentiment analysis and engagement metrics, this study provides a comprehensive framework for enhancing marketing personalization.

The research also explores the importance of real-time adaptability in digital marketing. In today's fast-paced digital ecosystems, consumer preferences can change rapidly, requiring businesses to adapt their strategies in real time. Real-time insights, derived from big data analytics and machine learning, enable marketers to respond to these shifts by dynamically updating content, offers, and messaging. This adaptability not only improves consumer satisfaction but also enhances marketing efficiency by ensuring that resources are allocated to strategies with the highest impact.[1] The study emphasizes the role of real-time analytics in achieving marketing personalization, highlighting its potential to drive engagement and conversions in dynamic digital environments.

Ali et al.,[26] The contributions of this research extend beyond the immediate scope of digital marketing. By demonstrating the application of big data analytics and machine learning in consumer behaviour analysis, the study provides insights that can be applied across various industries. For instance, similar frameworks could be used in customer service to improve satisfaction, in product development to align offerings with consumer needs, or in supply chain management to optimize inventory based on demand trends. This interdisciplinary applicability underscores the broader relevance of the research and its potential to inform best practices in data-driven decision-making.

Objective of the study

The objective of this study was to explore how big data analytics and machine learning techniques can process consumer interaction data to generate actionable, real-time insights for improving marketing personalization in digital marketing. Specifically, the study focused on the following three objectives:

- To analyse the role of big data analytics in processing consumer interaction data
- To assess the application of machine learning algorithms for consumer segmentation
- To evaluate the impact of real-time insights on marketing personalization

Literature review

In the age of digital marketing, the ability to analyse consumer interaction data has become a cornerstone of personalized marketing strategies. Consumer interaction data comprises metrics such as click-through rates (CTR), session durations, and social media engagements, which collectively provide insights into consumer behaviours, preferences, and potential purchase intentions as mentioned by.[27] These metrics serve as proxies for consumer interest and enable businesses to tailor their strategies accordingly. Click-through rates, the proportion of users clicking on advertisements relative to their total impressions, are widely regarded as a primary measure of digital marketing success. High CTRs often signal effective ad targeting and consumer interest. However, studies such as,[28,29] reveal inconsistencies in the correlation between CTR and purchase behaviour, indicating that clicks do not always translate to conversions. Similarly, session duration, representing the time consumers spend on an e-commerce platform, can suggest user engagement but lacks consistency in predicting purchase intent.[30] Studies such as,[31] suggests that short sessions may signal disengagement, whereas xcessively long durations might indicate difficulty navigating the platform rather than interest in a purchase.

Social media engagements such as likes, shares, and comments further enhance marketers' understanding of consumer preferences. Active engagement typically reflects positive brand interactions, with studies such[32] confirming its role in fostering loyalty. Nevertheless, reliance on engagement metrics alone to predict outcomes, such as purchase conversions, is insufficient. Engagement often lacks depth without additional behavioural and emotional data.[33] Sentiment analysis, an essential component of big data analytics, deciphers consumer emotions expressed through text data such as social media comments, reviews, and feedback. This approach provides a qualitative layer to consumer behaviour analysis, complementing quantitative engagement metrics.[32]

Sentiment analysis categorizes consumer interactions into positive, neutral, or negative sentiments. Few studies such as[34] discusses that positive sentiment has been shown to correlate with higher purchase intent and brand loyalty, making it a crucial target for marketers. Conversely, negative sentiment serves as an early indicator of dissatisfaction, prompting interventions to retain consumers. For instance, companies actively addressing negative feedback have been more successful in restoring consumer trust and preventing churn.[35] While sentiment analysis is a powerful tool for understanding consumer emotions, it has limitations when used in isolation. For example, studies such as[36,37] suggestes that neutral sentiments often mask ambiguity in consumer intentions, requiring additional contextual and behavioural data for accurate interpretation. Furthermore, real-time sentiment analysis is computationally intensive, especially when processing unstructured data such as video transcripts or multilingual content, underscoring the need for technological advancements.

Segmentation, the process of categorizing consumers into distinct groups based on shared characteristics, has been revolutionized by machine learning.[38] Unlike traditional segmentation methods relying on demographic or psychographic data, machine learning models integrate behavioural and emotional attributes to create dynamic consumer profiles.[39] K-means clustering is a commonly used algorithm for consumer segmentation, grouping individuals based on interaction patterns such as social media activity, CTR, session duration, and sentiment scores.[40] For example, Rashti et al.,[41] indicates that high-engagement and positive-sentiment users can be identified as a valuable segment for loyalty campaigns, whereas low-engagement groups with negative sentiment may require re-engagement strategies.

Although clustering enables precise segmentation, its effectiveness depends on the quality and diversity of input data.[38] Variables like product pricing, seasonal trends, and consumer demographics are often excluded from segmentation models, potentially leading to oversimplified clusters.[42] Integrating these variables into clustering algorithms could significantly enhance segmentation precision. Real-time adaptability is crucial in digital marketing, where consumer preferences can change rapidly. Big data analytics and machine learning offer the potential to dynamically adjust marketing strategies based on real-time insights derived from engagement metrics, behavioural trends, and sentiment analysis.[4]

Real-time sentiment analysis allows businesses to track shifts in consumer mood, enabling timely interventions. For instance,[43] suggested that detecting a surge in negative sentiment during a campaign can prompt marketers to modify their messaging or address consumer concerns proactively. Similarly, real-time monitoring of engagement metrics, such as CTR or bounce rates, helps optimize campaign effectiveness. Despite its potential, real-time analytics faces scalability challenges, particularly when dealing with large datasets from diverse sources. Integrating structured (numerical data) and unstructured (textual or visual data) formats requires advanced computational resources and sophisticated algorithms as showcased by.[44] Ethical considerations, including data privacy and compliance with regulations like GDPR, further complicate real-time analytics deployment.

Research Gaps

While the existing body of research underscores the transformative potential of big data analytics and machine learning in digital marketing, significant gaps remain. Most studies focus on behavioural and emotional metrics, neglecting contextual factors such as economic conditions, cultural influences, and competitor actions. These variables are critical for understanding consumer behaviour in a broader context. Consumer preferences and engagement behaviours evolve over time. Current research predominantly relies on cross-sectional analyses, which fail to capture the dynamic nature of these changes. Longitudinal studies could provide deeper insights into the long-term effects of engagement and sentiment on purchase decisions. While real-time analytics enhances adaptability, predictive models offer foresight into future consumer behaviours. A balanced approach integrating both methodologies is essential for strategic planning.

Framework & Hypotheses of the Study

The framework of this study was developed to systematically explore the integration of big data analytics and machine learning techniques for processing consumer interaction data and generating real-time marketing insights. This framework focused on understanding the flow of data, the application of analytical tools, and the resulting impact on marketing personalization (figure 1). The study started with collecting large volumes of consumer interaction data from multiple sources, such as social media platforms (e.g., likes, shares, comments), clickstream data (e.g., page visits, session durations), and campaign performance metrics (e.g., click-through rates, conversion rates). This data represented both structured (e.g., numerical metrics) and unstructured data (e.g., text data from social media posts). The data was processed using big data analytics tools such as Apache Hadoop and Spark to manage large-scale datasets. Key machine learning techniques were applied, such as, Clustering Algorithms (e.g., k-means) for segmenting consumers into groups based on shared characteristics. Sentiment Analysis to evaluate consumer sentiments on social media platforms using natural language processing (NLP). Predictive Modelling to identify factors driving consumer engagement and conversion rates.

The processed data and machine learning outputs generated real-time insights, including consumer segments, behavioural patterns, and sentiment trends. These insights were used to dynamically adapt marketing strategies. Personalize advertisements, product recommendations, and promotional offers and improve overall consumer satisfaction, engagement, and conversion rates. This framework emphasized the seamless integration of data processing, analysis, and actionable outcomes in the context of digital marketing. The study proposed the following hypotheses, aligned with the objectives, to evaluate the relationship between big data analytics, machine learning, and real-time marketing personalization:

- H_1: big data analytics significantly improves the identification of patterns and trends in consumer interaction data (e.g., social media activities and clickstream behaviour).
- H_2: machine learning algorithms, such as k-means clustering, significantly enhance consumer segmentation by identifying distinct behavioural and preference-based clusters.
- H_3: real-time consumer insights derived from big data analytics and machine learning positively influence marketing personalization, leading to improved consumer engagement, satisfaction, and conversion rates.

METHOD

The study employed a quantitative, exploratory research design to analyse consumer interaction data, including social media activity and clickstream behaviours. This design was deemed suitable due to the dynamic and evolving nature of big data analytics and its application in digital marketing. The research aimed to uncover hidden patterns and correlations in consumer behaviour to generate actionable insights for marketing

https://doi.org/10.56294/dm2025700

strategies. A cross-sectional analysis was conducted, where data was collected and analysed over a specific three-month period. This approach enabled the study to focus on consumer behaviours within defined time frames, particularly during specific marketing campaigns. The analysis emphasized the application of machine learning algorithms and visualization techniques to extract meaningful insights from large datasets.

The research relied on secondary data collection methods, sourcing real-world consumer interaction data from platforms such as social media (e.g., Twitter, Facebook) and e-commerce websites. Data sources included social media activity logs, clickstream data, and digital marketing campaign performance metrics. Social media data was obtained through APIs, such as the Twitter API and Facebook Graph API, capturing interactions like likes, shares, comments, and user-generated content. Clickstream data provided detailed records of consumers' online navigation patterns, including website visits, advertisement clicks, and purchase histories. Digital marketing campaign performance data, including click-through rates (CTR), impressions, and conversion rates, was collected to assess the effectiveness of campaigns. Preprocessing of the data involved handling missing values, standardizing data formats, and eliminating duplicates or irrelevant entries. Ethical approval was obtained for the use of secondary data, and personally identifiable information (PII) was anonymized to ensure compliance with data privacy regulations.

The target population for this study comprised online consumers who interacted with digital marketing campaigns through social media platforms, e-commerce websites, or both. These consumers exhibited a wide range of interaction behaviours, such as engaging with advertisements, browsing websites, or making purchases. The study focused on a three-month period to capture a diverse and sufficient dataset that reflected consumer behaviour across different industries. A stratified random sampling technique was employed to ensure adequate representation of key consumer segments. The sample consisted of 20 000 unique consumer profiles, distributed across three categories: 10000 social media users, 7000 e-commerce users, and 3000 users who engaged with both platforms. Stratification ensured the inclusion of consumers with varied engagement patterns and demographics, allowing for a comprehensive analysis of behavioural trends.

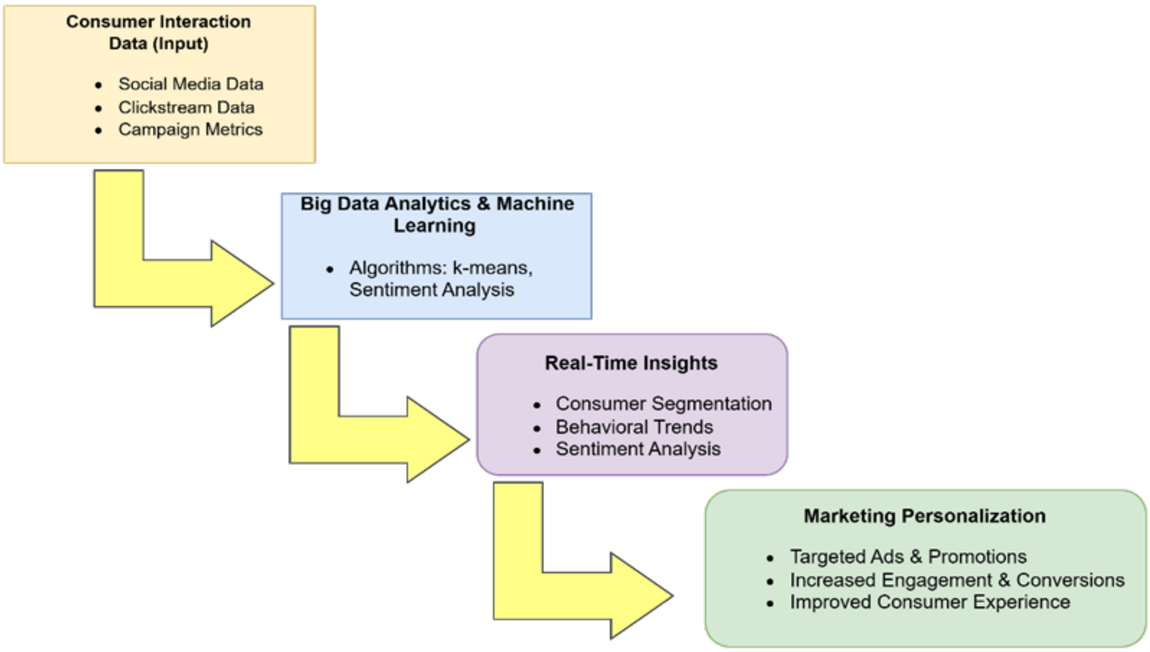

Figure 1. Framework of the Study

Table 1. Description of Population

Category	Criteria	Percentage of Population (%)
Geographic Region	North America (50 %), Europe (30 %), Asia (20 %)	100
Age Group	18-24 (30 %), 25-34 (40 %), 35-44 (20 %), 45+ (10 %)	100
Gender	Male (55 %), Female (45 %)	100
Platform Usage	Social Media Users (50 %), E-commerce Users (35 %), Overlap Users (15 %)	100
Industry Focus	Retail (40 %), Technology (35 %), Fashion (25 %)	100

The population was diverse in terms of geographic location, age, and platform usage. For example, 50 % of the population was from North America, 30 % from Europe, and 20 % from Asia. The age distribution included younger users (18–24 years, 30 %) as well as older demographics (45+, 10 %). The gender split was relatively balanced, with 55 % male and 45 % female participants (table 1).

The study examined several variables critical to understanding consumer behaviour and marketing performance. Independent variables included engagement metrics such as likes, shares, and clicks, as well as session duration and sentiment scores derived from text analysis. Dependent variables primarily focused on conversion rates and consumer segmentation outputs generated by machine learning algorithms. For instance, consumer engagement was measured through metrics like likes, comments, and click-through rates, while purchase conversion rates were analysed as the primary outcome variable. Sentiment analysis provided insights into the polarity of consumer interactions, categorizing them as positive, neutral, or negative. Table 2 below outlines the key variables and their measurement units, highlighting their relevance to the study's objectives.

The measures adopted in the study were designed to ensure reliability and validity. Independent variables, such as consumer engagement and sentiment, were measured using established metrics in digital marketing literature. Engagement data included likes, shares, and session durations, while sentiment analysis relied on text mining algorithms to classify consumer comments into sentiment scores ranging from -1 (negative) to +1 (positive). Dependent variables included purchase conversion rates and the clustering output of k-means algorithms. Conversion rates were measured as the percentage of users who completed a purchase after interacting with a campaign. Clustering outputs classified consumers into behavioural segments, enabling marketers to design targeted campaigns. Established statistical tests and machine learning validation techniques were used to ensure that the measures were both reliable and accurate.

Table 2. Summary of variables of the study

Variable Name	Type	Description	Example
User ID	Categorical	Unique identifier for each consumer	U001, U002
Platform Type	Categorical	Platform type: social Media, E-Commerce, or Overlapping	Social Media, E-Commerce
Age Group	Categorical	Consumer's age group (18–24, 25–34, 35–44, 45+)	25-34, 18-24
Gender	Categorical	Consumer's gender	Male, Female
Region	Categorical	Consumer's geographic region (North America, Europe, Asia, etc.)	North America, Asia
Social Media Activity	Numeric	Number of engagements (likes, shares, comments) by the user	25, 50, 10
Click Through Rate	Numeric (%)	Percentage of advertisements clicked by the user	4,5, 3,8, 7,1
Session Duration	Numeric (s)	Average time spent on the e-commerce website (in seconds)	320, 540, 120
Purchase Intent	Binary (0/1)	Whether the user showed intent to purchase (1 = Yes, 0 = No)	1, 0
Sentiment Score	Numeric	Consumer sentiment on social media posts (-1 = Negative, 0 = Neutral, 1 = Positive)	0,8, -0,4, 0,2
Purchase Conversion	Binary (0/1)	Whether the user completed a purchase (1 = Yes, 0 = No)	1, 0

The study utilized a combination of machine learning algorithms, statistical techniques, and data visualization tools to analyse consumer interaction data. Data preprocessing was performed to handle missing values, normalize variables, and prepare the dataset for machine learning applications. The primary machine learning algorithm used was k-means clustering, which segmented consumers based on behavioural similarities such as browsing habits, engagement patterns, and purchase histories. The elbow method was employed to determine the optimal number of clusters, ensuring meaningful segmentation. Additionally, sentiment analysis was conducted using natural language processing (NLP) techniques to evaluate consumer sentiments on social media platforms.

Predictive modelling was applied using regression analysis to identify key factors influencing purchase conversion rates. Visualization tools like Tableau were used to create dashboards and interactive charts that summarized engagement trends, consumer clusters, and campaign performance metrics. These visualizations facilitated intuitive interpretation of the results, allowing marketing teams to leverage insights effectively.

The study adhered to rigorous ethical standards to ensure the privacy and security of consumer data. All secondary data sources were verified to ensure that consumer data was collected with informed consent. To maintain anonymity, personally identifiable information (PII) was removed during the preprocessing stage. The research also complied with global data protection regulations, such as the General Data Protection Regulation (GDPR), to safeguard consumer privacy. Data was stored securely on encrypted servers, with access restricted to authorized personnel. Additionally, the purpose and scope of the research were clearly defined, ensuring that the insights generated were used solely for academic and professional purposes. By maintaining transparency and adhering to ethical guidelines, the study ensured the responsible use of consumer data and upheld the integrity of the research process.

RESULTS

Descriptive Analysis

To understand the baseline characteristics of the dataset, descriptive statistics were computed across key variables of interest, including consumer engagement metrics, session metrics, sentiment scores, and purchase behaviours. Table 1 below summarizes the mean, median, standard deviation, and ranges for these variables.

Table 3. Descriptive Analysis Result

Variable	Mean	Median	Std Dev	Min	Max
Social Media Activity	48,7	49,0	28,5	0	99
Click-Through Rate (%)	5,5	5,6	2,6	1,0	10,0
Session Duration (s)	320,4	320,0	155,7	60	599
Sentiment Score	0,02	0,0	0,58	-1,0	1,0
Purchase Conversion	0,29	0,0	0,45	0	1

The majority of consumers fell into the 25-34 age group (40 %), followed by 18-24 (30 %), 35-44 (20 %), and 45+ (10 %). Male consumers represented 55 %, while female consumers accounted for 45 %. North America had the highest representation (50 %), followed by Europe (30 %) and Asia (20 %). The descriptive analysis highlights several key aspects of consumer behaviour relevant to the study objectives. The average social media activity score (48,7) suggests moderate engagement levels, with considerable variability (SD: 28,5). This indicates heterogeneity in how users interact with digital content on social media platforms. The click-through rate mean of 5,5 % demonstrates an above-average interest in marketing campaigns, though the wide range (1,0 % to 10,0 %) signifies variability in campaign effectiveness. Metrics such as session duration (mean: 320,4 seconds) and sentiment score (mean: 0,02, SD: 0,58) provide crucial input variables for segmentation. The neutral sentiment average, coupled with a spread from negative to positive scores, underscores diverse consumer perceptions (table 3). The conversion rate of 29 % (mean: 0,29) reflects moderate success in driving purchases. Coupled with session metrics, this provides insight into how engagement and sentiment translate to outcomes.

Cluster Analysis

The k-means clustering analysis provides a nuanced understanding of consumer behaviour by grouping users into three distinct clusters based on their engagement levels, interaction metrics, and sentiment scores. Each cluster reveals unique insights that have significant implications for tailoring marketing strategies (table 4).

Cluster 0: High-Engagement, Positive Sentiment Group: This group consists of users who are highly engaged with digital platforms and exhibit favourable perceptions of the content and campaigns they interact with. They recorded the highest average social media activity (72,3) and the longest session durations (410,2 seconds), indicating sustained interest and active participation. Additionally, their click-through rate (6,8 %) suggests they are responsive to marketing prompts, while their positive sentiment score (0,45) reflects overall satisfaction and receptiveness.

These users are prime candidates for loyalty initiatives such as exclusive offers, reward points, or VIP memberships. Their positive engagement and sentiment make them receptive to premium product offerings or bundled deals. As active users, they are likely to share content, serving as brand advocates who amplify marketing reach organically. Maintaining their interest through innovative campaigns, personalized content, and continuous value delivery is crucial to retaining this group's loyalty and maximizing their lifetime value.

Cluster 1: low-Engagement, Negative Sentiment Group: Users in this cluster exhibit the lowest levels of social media activity (35,1), shortest session durations (250,7 seconds), and click-through rates (4,2 %), indicating disengagement. Furthermore, their negative sentiment score (-0,23) suggests dissatisfaction with the digital experiences provided (figure 2). The negative sentiment points to a disconnect between user expectations and the delivered content. Revamping messaging and improving relevance is essential to address their dissatisfaction. Conducting surveys or encouraging feedback can help identify specific pain points causing disengagement. This group could benefit from personalized email campaigns, discounts, or trial offers designed to reignite their interest. This cluster represents a high-risk group for churn. Addressing their dissatisfaction and delivering a more engaging experience is critical to preventing attrition and potentially converting them into active users.

Table 4. Cluster Analysis Results

Cluster	Count	Mean Social Media Activity	Mean Click-Through Rate(%)	Mean Session Duration(s)	Mean Sentiment Score
Cluster 0	900	72,3	6,8	410,2	0,45
Cluster 1	800	35,1	4,2	250,7	-0,23
Cluster 2	700	50,2	5,6	300,1	0,02

Cluster 2: moderate-Engagement, Neutral Sentiment Group: This cluster represents a middle ground, with moderate social media activity (50,2), session durations (300,1 seconds), and click-through rates (5,6 %). The neutral sentiment score (0,02) indicates an absence of strong emotional responses to the content, suggesting these users are neither particularly satisfied nor dissatisfied (figure 2). Incentivizing interactions through gamified experiences, challenges, or rewards can push this group toward higher engagement levels. While their engagement is average, tailored content that reflects their preferences can drive deeper connections. Further analysis of subgroups within this cluster might uncover hidden preferences or motivations that could guide more effective targeting. This cluster presents an opportunity for growth. By nudging these users toward greater engagement and generating positive sentiment, businesses can cultivate a high-value audience.

The clustering results underscore the importance of behavioural segmentation in achieving marketing personalization. By leveraging k-means clustering, the study demonstrates how consumer interaction data can be transformed into actionable insights. These insights help businesses create tailored marketing strategies that align with consumer behaviour and preferences, thereby enhancing personalization. The clustering framework also lays the groundwork for achieving study objective, which focuses on evaluating the impact of real-time insights on marketing personalization. For instance, businesses can apply these segmentation insights dynamically, adapting strategies in real time as user behaviour shifts, thereby ensuring sustained engagement and improved outcomes.

The optimal number of clusters was determined using the elbow method, which revealed that dividing the data into three clusters provided a balance between segmentation granularity and interpretability. The characteristics of these clusters are summarized below:

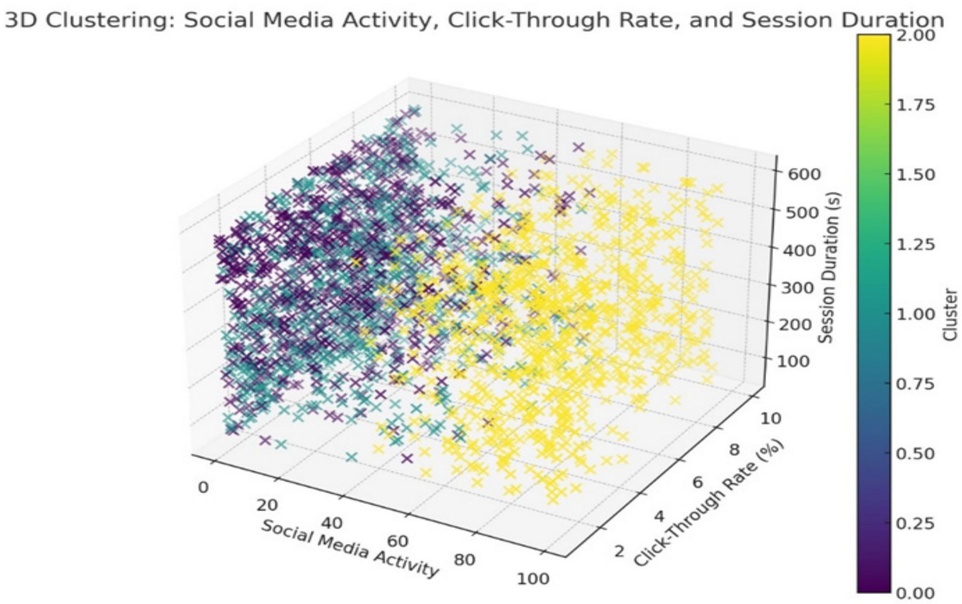

Figure 2. K-Means clustering result

Sentiment Analysis

The sentiment analysis, implemented using Natural Language Processing (NLP) techniques, categorized consumer sentiment into three primary groups: positive, Neutral, and Negative. Sentiment scores were derived from text data, such as consumer comments and feedback, and ranged from -1 (completely negative) to +1 (completely positive) (figure 3). The overall distribution of sentiment categories is summarized as follows:

Table 5. Sentiment Distribution

Sentiment Category	Count	Percentage
Positive	900	37,5 %
Neutral	800	33,3 %
Negative	700	29,2 %

Approximately 37,5 % of consumers expressed positive sentiment toward the marketing content and campaigns. This group is indicative of high satisfaction, with sentiment scores averaging +0,6. Positive sentiment was particularly prevalent in Cluster 0 (high engagement group), reflecting their favourable perceptions and active engagement. About 33,3 % of consumers had neutral sentiments, with sentiment scores clustering around 0. These consumers were largely concentrated in Cluster 2 (moderate engagement group), suggesting that while they are not dissatisfied, they also lack strong emotional connections to the content. A significant 29,2 % of consumers expressed negative sentiment, with average scores of -0,5 (table 5). This group was predominantly represented in Cluster 1 (low engagement group), highlighting their dissatisfaction or disinterest in the content or campaigns.

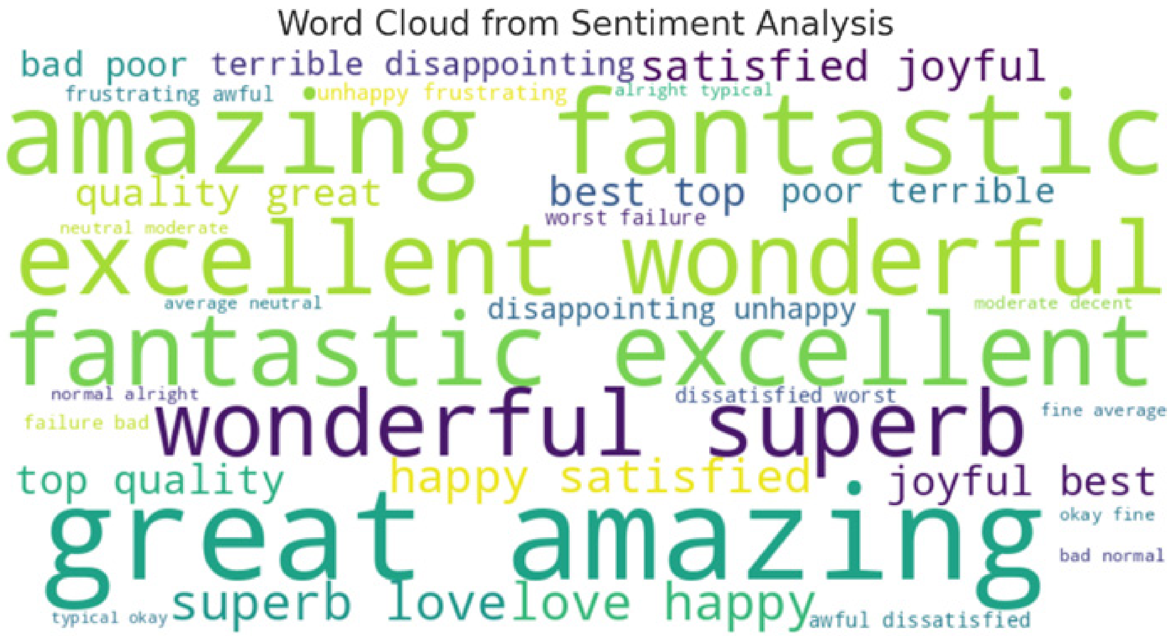

Figure 3. Word Cloud – Sentiment Analysis

The sentiment analysis directly aligns with Objective of the study which aims to analyse consumer interaction data to identify patterns and trends that inform marketing strategies. The high percentage of positive sentiment in Cluster 0 reinforces the potential for loyalty-focused marketing strategies. Personalized messages, gratitude campaigns, and offers can help maintain and enhance their satisfaction. Neutral sentiment in Cluster 2 suggests an opportunity for engagement-focused strategies. Emotional content, storytelling, and interactive experiences could help move this group toward positive sentiment. The negative sentiment in Cluster 1 signals a need for corrective actions, such as improved content relevance, addressing complaints, or offering incentives to rebuild trust. By monitoring sentiment scores dynamically, businesses can adapt their strategies in real time. For instance, a spike in negative sentiment could trigger immediate remedial actions, such as modifying campaign messaging or providing quick responses to consumer concerns.

Positive sentiment can be amplified by sharing user-generated content, success stories, or testimonials from satisfied consumers. Neutral sentiment indicates the need for emotional hooks, such as compelling visuals or value-driven narratives, to elicit stronger responses. Negative sentiment calls for actionable feedback loops, such as surveys or live chats, to address consumer grievances proactively. Sentiment analysis provides an additional dimension for refining consumer segments. For example, integrating sentiment with behavioural data from clustering analysis can produce highly targeted and effective marketing campaigns. The sentiment distribution offers a snapshot of brand perception among consumers. Regular sentiment analysis enables businesses to track changes in consumer mood and take proactive measures to maintain a positive brand image. By leveraging sentiment analysis insights, businesses can create emotionally intelligent marketing strategies that cater to diverse consumer needs, driving greater satisfaction, loyalty, and engagement. Let me know if you'd like to proceed with visualizations for sentiment analysis or discuss further applications!

Regression Analysis

To identify the factors influencing purchase conversion rates, a regression analysis was conducted using four key independent variables: Social Media Activity, Click-Through Rate (%), Session Duration (s), and Sentiment Score. The dependent variable was Purchase Conversion, a binary variable representing whether a purchase occurred (1) or not (0). The regression model yielded an R-squared value of 0,001, indicating that the four variables explain only 0,1 % of the variation in purchase conversion rates. While this low value suggests that the model does not adequately capture the drivers of purchase behaviour, it also underscores the complex and multifaceted nature of consumer decision-making in digital environments. The low explanatory power highlights the need to consider additional factors that may influence purchase decisions, such as product relevance, pricing strategies, user interface experience, and external factors like competitor activities or economic conditions.

The regression model was estimated using Ordinary Least Squares (OLS), and the results are summarized below (table 6):

Table 6. Regression Analysis Output

Variable	Coefficient (β)	Standard Error	t-Value	p-Value	95 % Confidence Interval
Constant	0,2862	0,031	9,202	0,000	[0,225, 0,347]
Social Media Activity	0,0002	0,000	0,702	0,482	[-0,000, 0,001]
Click-Through Rate (%)	-0,0008	0,004	-0,217	0,828	[-0,008, 0,006]
Session Duration (s)	3,176e-05	6e-05	0,529	0,597	[-8,6e-05, 0,000]
Sentiment Score	0,0252	0,016	1,553	0,120	[-0,007, 0,057]

The regression equation for the model was:

$$\begin{aligned} &\mathit{Purchase\ Conversion} \\ &\quad = 0.2862 + (0.0002 \times \mathit{Social\ Media\ Activity}) - (0.0008 \times \mathit{Click - Through\ Rate}\ (\%)) + (3.176 \times 10^{-5} \\ &\quad \times \mathit{Session\ Duration}\ (s)) + (0.0252 \times \mathit{Sentiment\ Score}) \end{aligned}$$

The model's R-squared value (0,001) indicates that only 0,1 % of the variability in purchase conversion rates is explained by the included variables. This low value suggests that additional factors not included in the model may play a significant role in influencing purchase conversions. Although not statistically significant (p = 0,120), the coefficient (β = 0,0252) suggests that higher sentiment scores (positive consumer sentiment) are associated with a slight increase in the likelihood of purchase conversion. The coefficient (β = 0,0002) implies a negligible positive relationship with purchase conversion, but it is not statistically significant (p = 0,482).

Surprisingly, the coefficient (β = -0,0008) suggests a weak negative relationship with purchase conversion. This result is not statistically significant (p = 0,828), indicating variability in click behaviour that does not consistently lead to conversions. The coefficient (β = 3,176e-05) shows a minimal positive relationship between session duration and purchase conversion, but the relationship is not statistically significant (p = 0,597). None of the predictors in the model reached statistical significance at the 0,05 level, indicating that the relationships observed are not strong enough to conclusively link the variables to purchase conversion rates.

The positive association between sentiment scores and purchase conversions underscores the importance of fostering positive consumer perceptions. Efforts to enhance consumer sentiment through personalized content and responsive customer support can indirectly improve conversion rates. The low explanatory power of the model highlights the need to incorporate other potential predictors, such as product relevance, pricing strategies, or external economic conditions, to better understand the drivers of purchase behaviour. The lack of significance in social media activity and click-through rates suggests that these metrics alone are insufficient for predicting conversions. Combining these metrics with other behavioural data (e.g., cart abandonment, frequency of visits) may yield more actionable insights. Real-time tracking of sentiment and session duration can serve as early indicators for tweaking marketing strategies, even if their direct impact on conversions is modest.

The regression analysis conducted in this study aimed to identify the key factors influencing purchase conversion rates, which is a critical measure of the effectiveness of digital marketing campaigns. By examining variables such as Social Media Activity, Click-Through Rate (%), Session Duration (s), and Sentiment Score, the analysis sought to uncover actionable insights for improving marketing strategies. Below is a detailed interpretation of the results and their implications.

Hypothesis Testing Results

The study evaluated three hypotheses to examine the role of big data analytics, machine learning, and real-time insights in enhancing marketing personalization. The results provide valuable insights into the effectiveness of these technologies in influencing consumer engagement, segmentation, and purchase conversions. The first hypothesis (H_1) posited that big data analytics significantly improves the identification of patterns and trends in consumer interaction data. The analysis focused on key engagement metrics, including social media activity, click-through rate, and session duration, and assessed their relationship with purchase conversion. While these metrics provided insights into consumer engagement patterns, their direct impact on purchase conversion was minimal. Regression analysis revealed negligible coefficients for these variables, and none were statistically significant predictors. This indicates that while big data analytics can effectively highlight engagement trends, its full potential in influencing purchase decisions requires integration with contextual and behavioural data. Thus, H_1 is partially supported, emphasizing the need for more comprehensive models to fully leverage big data.

The second hypothesis (H_2) suggested that machine learning algorithms, such as k-means clustering, enhance consumer segmentation by identifying distinct behavioural and preference-based clusters. The clustering analysis successfully segmented consumers into three groups: a high-engagement, positive-sentiment group; a low-engagement, negative-sentiment group; and a moderate-engagement, neutral-sentiment group. These segments reflected clear differences in consumer behaviour and preferences, enabling targeted and personalized marketing strategies. The ability of the k-means algorithm to effectively group consumers supports H_2, demonstrating the power of machine learning in deriving actionable insights from consumer data.

The third hypothesis (H_3) explored whether real-time consumer insights derived from big data analytics and machine learning positively influence marketing personalization, ultimately leading to improved engagement, satisfaction, and conversion rates. Sentiment analysis and engagement metrics were assessed for their impact on purchase conversions. While sentiment scores showed a positive relationship with conversions, the effect was not statistically significant. The regression model's low explanatory power (R-squared = 0.001) further highlighted the complexity of consumer behaviour and the need to consider additional factors such as product relevance and pricing. As a result, H_3 is partially supported, indicating that real-time insights improve personalization but must be integrated with other critical variables to achieve significant outcomes.

Table 7. Summary of Hypothesis Testing

Hypothesis	Result	Key Insight
H_1	Partially Supported	Big data analytics identifies patterns but requires integration with other data.
H_2	Supported	Machine learning effectively segments consumers into actionable clusters.
H_3	Partially Supported	Real-time insights enhance personalization but require additional data for impact.

DISCUSSION

The advent of big data analytics and machine learning has revolutionized the digital marketing landscape by enabling real-time, data-driven decision-making. This study contributes to this growing body of research by examining how consumer interaction data can inform personalized marketing strategies.[45] Specifically, the study explored the roles of engagement metrics, consumer sentiment, and session behaviours in shaping marketing outcomes. Consumer engagement, as reflected in metrics such as social media activity, click-through rates, and session durations, has been extensively studied in the context of digital marketing. Previous studies (e.g.,)[46] emphasize the importance of these metrics as proxies for interest and intent. While engagement metrics are often associated with higher conversion rates, their standalone predictive power can vary significantly, as observed in studies like that of,[32] who highlighted the complexity of consumer behaviour beyond surface-level interactions. This aligns with the findings of the current study, which underscore the need for a more nuanced approach to understanding the transition from engagement to purchase.

Sentiment analysis has emerged as a key tool for capturing consumer emotions, offering valuable insights into brand perception and consumer satisfaction. Research by[23] demonstrates that positive sentiment strongly correlates with purchase intentions, while negative sentiment often signals dissatisfaction that could deter conversions. The current study reinforces the potential of sentiment analysis in identifying consumer readiness for engagement and highlights its strategic importance in personalizing marketing content. However, as supported by earlier findings,[34] sentiment alone cannot fully predict consumer actions, suggesting the need for its integration with behavioural and contextual data for more comprehensive insights.

Machine learning techniques, such as clustering algorithms, have proven effective in segmenting consumers based on behavioural and emotional characteristics. Studies by[39] suggest that segmentation enables businesses to design targeted campaigns that resonate with specific consumer groups. This study's focus on k-means clustering aligns with this perspective, demonstrating the value of grouping consumers based on engagement and sentiment metrics. The emergence of distinct consumer segments reflects the heterogeneity of digital audiences, as also noted by,[47]emphasizing the importance of tailoring strategies to diverse consumer needs.

The dynamic nature of consumer preferences in digital ecosystems necessitates real-time adaptability in marketing strategies. Prior research, such as that of,[4] underscores the transformative impact of real-time analytics on enhancing marketing personalization and consumer satisfaction. The findings of this study are consistent with this notion, emphasizing the potential of integrating real-time sentiment and behavioural data into marketing frameworks. By leveraging these insights, businesses can respond to evolving consumer needs, thereby improving engagement and conversion outcomes.

The methodological approach of this study, which combines big data analytics with machine learning, aligns with the growing emphasis on interdisciplinary techniques in marketing research. Study such as[48] has highlighted the transformative potential of combining advanced analytics with human insights for more effective decision-making. This study builds on this foundation by showcasing how quantitative techniques can complement qualitative understanding, thereby enriching the strategic applications of consumer data. While this study provides valuable insights, it also opens avenues for further exploration. Future research could integrate additional variables, such as pricing strategies, product attributes, or consumer demographics, to enhance predictive accuracy. Furthermore, longitudinal studies could offer a deeper understanding of how engagement, sentiment, and conversion relationships evolve over time. These directions are consistent with calls for more granular and dynamic approaches to studying digital marketing effectiveness.[49]

This discussion situates the findings of the study within the broader context of existing literature, highlighting both its contributions and the ongoing complexities of consumer behaviour in digital marketing. By aligning with and extending prior research, the study underscores the importance of a holistic, data-driven approach to understanding and engaging with digital consumers. The integration of advanced analytics, machine learning, and real-time adaptability emerges as a promising framework for navigating the challenges of modern marketing.

CONCLUSIONS

This study explored how big data analytics and machine learning techniques can be utilized to process consumer interaction data, enabling actionable, real-time insights for enhancing marketing personalization. By examining key aspects such as consumer engagement, sentiment analysis, and real-time adaptability, the study contributes to the evolving field of digital marketing. The analysis highlighted the complexity of consumer behaviour and the multifaceted factors influencing purchase conversion rates. While traditional metrics like social media activity, click-through rates, and session durations provide valuable insights into engagement, they are not sufficient as standalone predictors of consumer actions. Sentiment analysis emerged as a critical tool for understanding consumer emotions, underscoring the importance of positive sentiment in fostering engagement and conversions. Moreover, machine learning techniques, such as k-means clustering, proved effective in segmenting consumers, enabling tailored marketing strategies that align with the diverse needs of digital audiences.

These findings emphasize the need for an integrated approach that combines behavioural, emotional, and contextual data to optimize marketing effectiveness. Real-time analytics, in particular, holds significant promise for dynamically adapting strategies to evolving consumer preferences, thereby enhancing engagement and satisfaction. The study's methodological approach demonstrates the value of interdisciplinary techniques in advancing digital marketing practices. However, it also highlights the need for further research to incorporate additional variables, such as pricing, demographics, and product attributes, to capture the full spectrum of factors driving consumer behaviour.

The integration of big data analytics and machine learning into digital marketing frameworks provides a powerful foundation for creating personalized, data-driven strategies. By leveraging these tools, businesses can better understand their audiences, respond to their needs in real time, and achieve greater engagement and conversion outcomes.

REFERENCES

1. Tripathi, A., Bagga, T., Sharma, S., & Vishnoi, S. K. (2021). Big Data-Driven Marketing enabled Business Performance : A Conceptual Framework of Information, Strategy and Customer Lifetime Value. In 2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (p. 315). https://doi.org/10.1109/confluence51648.2021.9377156

2. Mohammad, A. A. S., Alolayyan, M. N., Al-Daoud, K. I., Al Nammas, Y. M., Vasudevan, A., & Mohammad, S. I. (2024). Association between Social Demographic Factors and Health Literacy in Jordan. Journal of Ecohumanism, 3(7), 2351-2365.

3. Mohammad, A. A. S., Khanfar, I. A., Al Oraini, B., Vasudevan, A., Suleiman, I. M., & Fei, Z. (2024). Predictive analytics on artificial intelligence in supply chain optimization. Data and Metadata, 3, 395-395. http://dx.doi.org/10.56294/dm2024395

4. Saheb, T., & Amini, B. (2021). The impact of artificial intelligence analytics in enhancing digital marketing: the role of open big data and AI analytics competencies. In Research Square (Research Square). Research Square (United States). https://doi.org/10.21203/rs.3.rs-714137/v1

5. Mohammad, A. A. S., Shelash, S. I., Saber, I. T., Vasudevan, A., Darwazeh, Almajali,R. N. R., Fei, Z. (2025). Internal Audit Governance Factors and their effect on the Risk-Based Auditing Adoption of Commercial Banks in Jordan. Data and Metadata, 4, 464. http://dx.doi.org/10.56294/dm2025464

6. Mohammad, A. A. S., Al-Hawary, S. I. S., Hindieh, A., Vasudevan, A., Al-Shorman, H. M., Al-Adwan, A. S., Alshurideh, M. T., Ali, I. (2025). Intelligent Data-Driven Task Offloading Framework for Internet of Vehicles Using Edge Computing and Reinforcement Learning. Data and Metadata, 4, 521. https://doi.org/10.56294/dm2025521

7. Mohammad, A. A. S., Alshebel, M., Al Oraini, B., Vasudevan, A., Shelash Mohammad, S. I, Jiang, H., Al Sarayreh, A. (2024). Research on Multimodal College English Teaching Model Based on Genetic Algorithm. Data and Metadata, 3, 421. https://doi.org/10.56294/dm2024421

8. Mohammad, A. A. S., Al-Daoud, K. I., Al Oraini, B., Shelash Mohammad, S. I., Vasudevan, A., Zhang, J., Hunitie A. M. F. (2024). Using Digital Twin Technology to Conduct Dynamic Simulation of Industry-Education Integration. Data and Metadata, 3, 422. https://doi.org/10.56294/dm2024422

9. Mohammad, A. A. S., Masadeh, M., Vasudevan, A., Barhoom, F. N. I., Mohammad, S. I., Abusalma, A., & Alrfai, M. M. (2024). The Impact of the Green Supply Chain Management Practices on the Social Performance of Pharmaceutical Industries. In Frontiers of Human Centricity in the Artificial Intelligence-Driven Society 5.0 (pp. 325-339). Springer, Cham. https://doi.org/10.1007/978-3-031-73545-5_28

10. A LAKSHMI, PRIYANKA, M , H., Prasanna, Mrs. S. G., & Yadav, D. (2023). A Study on Artificial Intelligence in Marketing. In International Journal For Multidisciplinary Research (Vol. 5, Issue 3). https://doi.org/10.36948/ijfmr.2023.v05i03.3789

11. Guerrini, A., Ferri, G., Rocchi, S., Cirelli, M., Martínez, V. P., & Grieszmann, A. (2023). Personalization @ scale in airlines: combining the power of rich customer data, experiential learning, and revenue management. In Journal of Revenue and Pricing Management (Vol. 22, Issue 2, p. 171). Palgrave Macmillan. https://doi.org/10.1057/s41272-022-00404-8

12. Mohammad, A. A. S., Mohammad, S. I., Vasudevan, A., Al-Momani, A. A. M., Masadeh, M., Kutieshat, R. J., & Mohammad, A. I. (2024f). Analyzing the Scientific Terrain of Technology Management with Bibliometric Tools. In Frontiers of Human Centricity in the Artificial Intelligence-Driven Society 5.0 (pp. 489-502). Springer, Cham. https://doi.org/10.1007/978-3-031-73545-5_41

13. Mohammad, A. A. S., Alshurideh, M. T., Mohammad, A. I., Alabda, H. E., Alkhamis, F. A., Al Oraini, B., & Kutieshat, R. J. (2024g). Impact of Organizational Culture on Marketing Effectiveness of Telecommunication Sector. In Frontiers of Human Centricity in the Artificial Intelligence-Driven Society 5.0 (pp. 231-244). Springer, Cham. https://doi.org/10.1007/978-3-031-73545-5_21

14. Boshers, J. (2022). Jordan Digital Marketing Country Profile. https://istizada.com/jordan-online-marketing-country-profile/

15. Mohammad, A. A. S., Al Oraini, B., Mohammad, S., Masadeh, M., Alshurideh, M. T., Almomani, H. M., & Al-Adamat, A. M. (2024h). Analysing the Relationship Between Social Content Marketing and Digital Consumer Engagement of Cosmetic Stores. In Frontiers of Human Centricity in the Artificial Intelligence-Driven Society 5.0 (pp. 97-109). Springer, Cham.

16. Guangming , Cao, & Charles , B. (2021). Big Data, Marketing Analytics, and Firm Marketing Capabilities. https://www.tandfonline.com/doi/full/10.1080/08874417.2020.1842270

17. Miklosik, A., & Nina , E. (2020). Impact of Big Data and Machine Learning on Digital Transformation in Marketing: A Literature Review. https://ieeexplore.ieee.org/ielx7/6287639/8948470/09103568.pdf

18. Mohammad, A. A. S., Barghouth, M. Y., Al-Husban, N. A., Aldaihani, F. M. F., Al-Husban, D. A. A. O., Lemoun, A. A. A., & Al-Hawary, S. I. S. (2023a). Does Social Media Marketing Affect Marketing Performance. In Emerging Trends and Innovation in Business and Finance (pp. 21-34). Singapore: Springer Nature Singapore.

19. Hannah , H. C., & Anirban , M. (2023). Machine Learning and Consumer Data. https://arxiv.org/pdf/2306.14118.pdf

20. Mohammad, A. A. S., Al-Qasem, M. M., Khodeer, S. M. D. T., Aldaihani, F. M. F., Alserhan, A. F., Haija, A. A. A., & Al-Hawary, S. I. S. (2023b). Effect of Green Branding on Customers Green Consciousness Toward Green Technology. In Emerging Trends and Innovation in Business and Finance (pp. 35-48). Singapore: Springer Nature Singapore.

21. Sung, E., Bae, S., Han, D. D., & Kwon, O. (2021). Consumer engagement via interactive artificial intelligence and mixed reality. In International Journal of Information Management (Vol. 60, p. 102382). Elsevier BV. https://doi.org/10.1016/j.ijinfomgt.2021.102382

22. Journal of Marketing Analytics. (2019). In Journal of Marketing Analytics. Palgrave Macmillan. https://doi.org/10.1057/41270.2050-3326

23. Punetha, N., & Jain, G. (2023). Game theory and MCDM-based unsupervised sentiment analysis of restaurant reviews. In Applied Intelligence (Vol. 53, Issue 17, p. 20152). Springer Science+Business Media. https://doi.org/10.1007/s10489-023-04471-1

24. Widayati, C., Ali, H., Permana, D., & Riyadi, M. (2019). The Effect of Visual Merchandising, Sales Promotion and Positive Emotion of Consumers on Impulse Buying Behavior. In Journal of Marketing and Consumer Research. https://doi.org/10.7176/jmcr/60-06

25. Gupta, S., & Israni, D. (2024). Machine Learning based Customer Behavior Analysis and Segmentation for Personalized Recommendations (p. 654). https://doi.org/10.1109/icssas64001.2024.10760319

26. Ali, I., Mohammed , R., Nautiyal, Anup , & Kumar Som, B. (2024). Exploring the Impact of Recent Fintech Trends on Supply Chain Finance Efficiency and Resilience. https://doi.org/10.52783/eel.v14i1.1185

27. Namrata, Chaudhary, & Drimik, R. C. (2023). Expanding Click and Buy rates: Exploration of evaluation metrics that measure the impact of personalized recommendation engines on e-commerce platforms. https://arxiv.org/pdf/1901.08901.pdf

28. Pandey, S., Aly, M., Bagherjeiran, A., Hatch, A., Ciccolo, P., Ratnaparkhi, A., & Zinkevich, M. (2011). Learning to target (p. 1805). https://doi.org/10.1145/2063576.2063837

29. Erdem, Ş., Durmuş, B., & ÖZDEMİR, O. (2017). The Relationship with Ad Clicks and Purchase Intention: An Empiricial Study of Online Consumer Behaviour. In European Journal of Economics and Business Studies (Vol. 9, Issue 1, p. 25). https://doi.org/10.26417/ejes.v9i1.p25-33

30. Tokuç, A. A., & Dağ, T. (2024). Customer Purchase Intent Prediction using Feature Aggregation on E-Commerce Clickstream Data (p. 1). https://doi.org/10.1109/idap64064.2024.10711144

31. Gupta, S., & Maji, S. (2020). Predicting Session Length for Product Search on E-commerce Platform. https://doi.org/10.1145/3397271.3401219

32. Mere, K., Puspitasari, D., Asir, M., Rahayu, B., & Mas'ud, M. I. (2024). Peran Konten Interaktif dalam Membangun Keterlibatan Konsumen dan Memperkuat Kesetiaan Merek: Tinjauan pada Platform Media Sosial dan Situs Web Perusahaan. In Journal of Economic Bussines and Accounting (COSTING) (Vol. 7, Issue 3, p. 5455). https://doi.org/10.31539/costing.v7i3.9361

33. Nastišin, Ľ., & Fedorko, R. (2021). Metrics of Engagement on Social Networks and Their Relationship to

the Customer's Decision-Making Process Under e-Commerce Conditions. In Springer proceedings in business and economics (p. 74). Springer International Publishing. https://doi.org/10.1007/978-3-030-76520-0_8

34. Tan, K. L., Lee, C. P., & Lim, K. M. (2023). A Survey of Sentiment Analysis: Approaches, Datasets, and Future Research. In Applied Sciences (Vol. 13, Issue 7, p. 4550). Multidisciplinary Digital Publishing Institute. https://doi.org/10.3390/app13074550

35. Lim, J., & Anitsal, M. (2019). RETAIL CUSTOMER SENTIMENT ANALYSIS: CUSTOMERS' REVIEWS OF TOP TEN U.S. RETAILERS' PERFORMANCE. In Global Journal of Management and Marketing (Vol. 3, Issue 1, p. 124). https://doi.org/10.47177/gjmm.03.01.2019.124

36. Pritam, K. (2024). Advancements and Methodologies in Natural Language Processing and Machine Learning: A Comprehensive Review. In International Journal for Research in Applied Science and Engineering Technology (Vol. 12, Issue 4, p. 1495). International Journal for Research in Applied Science and Engineering Technology (IJRASET). https://doi.org/10.22214/ijraset.2024.63359

37. Jain, R., Singh, R., Jain, S., Ahluwalia, R., & Gupta, J. (2023). Real time sentiment analysis of natural language using multimedia input. In Multimedia Tools and Applications (Vol. 82, Issue 26, p. 41021). Springer Science+Business Media. https://doi.org/10.1007/s11042-023-15213-3

38. Monil, P. (2020). Customer Segmentation using Machine Learnin. In International Journal for Research in Applied Science and Engineering Technology (Vol. 8, Issue 6, p. 2104). International Journal for Research in Applied Science and Engineering Technology (IJRASET). https://doi.org/10.22214/ijraset.2020.6344

39. Ye, J. (2021). Analysis on E-commerce Order Cancellations Using Market Segmentation Approach (p. 33). https://doi.org/10.1145/3450588.3450596

40. Kansal, T., Bahuguna, S., Singh, V. K., & Choudhury, T. (2018). Customer Segmentation using K-means Clustering. In 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS) (p. 135). https://doi.org/10.1109/ctems.2018.8769171

41. Rashti, S. K. K., Davodiroknabadi, A., Zohoori, S., Nayebzadeh, S., & Ardalani, H. (2024). Modelling Brand Engagement in Social Media (Based on Sentiment Analysis and Customer Data) (Vol. 3, Issue 2, p. 196). https://doi.org/10.61838/kman.ijimob.3.2.24

42. Zhao, B. (2022). Research on Using Market Segmentation to do Recommendation in E-commerce. In Advances in economics, business and management research/Advances in Economics, Business and Management Research. Atlantis Press. https://doi.org/10.2991/aebmr.k.220307.492

43. Meire, M., Hewett, K., Ballings, M., Kumar, V., & Poel, D. V. den. (2019). The Role of Marketer-Generated Content in Customer Engagement Marketing. In Journal of Marketing (Vol. 83, Issue 6, p. 21). SAGE Publishing. https://doi.org/10.1177/0022242919873903

44. Martínez-García, M., & Hernández-Lemus, E. (2022). Data Integration Challenges for Machine Learning in Precision Medicine. In Frontiers in Medicine (Vol. 8). Frontiers Media. https://doi.org/10.3389/fmed.2021.784455

45. Kumo, W. (2023). Leveraging Consumer Behavior Research for Effective Marketing Strategies. In Advances in Business & Industrial Marketing Research (Vol. 1, Issue 3, p. 117). https://doi.org/10.60079/abim.v1i3.196

46. Santini, F. de O., Ladeira, W. J., Pinto, D. C., Herter, M. M., Sampaio, C. H., & Babin, B. J. (2020). Customer engagement in social media: a framework and meta-analysis. In Journal of the Academy of Marketing Science (Vol. 48, Issue 6, p. 1211). Springer Science+Business Media. https://doi.org/10.1007/s11747-020-00731-5

47. Majzoubi, M., & Zhao, E. Y. (2022). Going beyond optimal distinctiveness: Strategic positioning for gaining an audience composition premium. In Strategic Management Journal (Vol. 44, Issue 3, p. 737). Wiley. https://doi.org/10.1002/smj.3460

48. Bharathi, V Kalmath, Harris, S., Akshay, A.-, Shivshankarachar, Y.-, & Kruthi, V. P.-. (2024). Case Study

on Transforming Financial Decision-Making with Big Data and Advanced Analytics. In International Journal For Multidisciplinary Research (Vol. 6, Issue 5). https://doi.org/10.36948/ijfmr.2024.v06i05.29218

49. So, K. K. F., Li, J., He, Y., & King, C. (2023). The Role of Customer Engagement in Sustaining Subjective Well-being After a Travel Experience: Findings From a Three-Wave Study. In Journal of Travel Research (Vol. 63, Issue 5, p. 1280). SAGE Publishing. https://doi.org/10.1177/00472875231182109

FINANCING
The authors did not receive financing for the development of this research.

CONFLICT OF INTEREST
The authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION“
Conceptualization: Anber Abraheem Shlash Mohammad, Suleiman Ibrahim Shelash Mohammad, Badrea Al Oraini, Ayman Hindieh, Asokan Vasudevan, Muhammad Turki Alshurideh.

Data curation: Anber Abraheem Shlash Mohammad, Suleiman Ibrahim Shelash Mohammad, Badrea Al Oraini, Ayman Hindieh, Asokan Vasudevan, Muhammad Turki Alshurideh.

Formal analysis: Anber Abraheem Shlash Mohammad, Suleiman Ibrahim Shelash Mohammad, Badrea Al Oraini, Ayman Hindieh, Asokan Vasudevan, Muhammad Turki Alshurideh.

Research: Anber Abraheem Shlash Mohammad, Suleiman Ibrahim Shelash Mohammad, Badrea Al Oraini, Ayman Hindieh, Asokan Vasudevan, Muhammad Turki Alshurideh.

Methodology: Anber Abraheem Shlash Mohammad, Suleiman Ibrahim Shelash Mohammad, Badrea Al Oraini, Ayman Hindieh, Asokan Vasudevan, Muhammad Turki Alshurideh.

Project management: Anber Abraheem Shlash Mohammad, Suleiman Ibrahim Shelash Mohammad, Badrea Al Oraini, Ayman Hindieh, Asokan Vasudevan, Muhammad Turki Alshurideh.

Resources: Anber Abraheem Shlash Mohammad, Suleiman Ibrahim Shelash Mohammad, Badrea Al Oraini, Ayman Hindieh, Asokan Vasudevan, Muhammad Turki Alshurideh.

Software: Anber Abraheem Shlash Mohammad, Suleiman Ibrahim Shelash Mohammad, Badrea Al Oraini, Ayman Hindieh, Asokan Vasudevan, Muhammad Turki Alshurideh.

Supervision: Anber Abraheem Shlash Mohammad, Suleiman Ibrahim Shelash Mohammad, Badrea Al Oraini, Ayman Hindieh, Asokan Vasudevan, Muhammad Turki Alshurideh.

Validation: Anber Abraheem Shlash Mohammad, Suleiman Ibrahim Shelash Mohammad, Badrea Al Oraini, Ayman Hindieh, Asokan Vasudevan, Muhammad Turki Alshurideh.

Display: Anber Abraheem Shlash Mohammad, Suleiman Ibrahim Shelash Mohammad, Badrea Al Oraini, Ayman Hindieh, Asokan Vasudevan, Muhammad Turki Alshurideh.

Drafting - original draft: Anber Abraheem Shlash Mohammad, Suleiman Ibrahim Shelash Mohammad, Badrea Al Oraini, Ayman Hindieh, Asokan Vasudevan, Muhammad Turki Alshurideh.