Advances in Consumer Research
https://acr-journal.com/
Volume-2 | Issue-5 | Nov 2025
Original Researcher Article

Multimodal Big Data Analytics for Customer Journey Optimization Across Digital Platforms

Dr. Praveen Kumar
Assistant Professor, GGSIPU, New Delhi, drpraveenkumar11@outlook.com, 0000-0001-6789-4828

Received: 10/10/2025
Revised: 17/10/2025
Accepted:18/11/2025
Published:28/11/2025

ABSTRACT

Contemporary customers interact with brands through dense, heterogeneous streams of data generated across websites, mobile applications, social media, conversational interfaces, and physical touchpoints. These interactions form complex, nonlinear customer journeys that can no longer be adequately understood through channel-specific or unimodal analytics. This paper proposes a comprehensive framework for multimodal big data analytics aimed at optimizing customer journeys across digital platforms. The framework integrates structured transactional logs, clickstream sequences, textual reviews, visual content, and interactional signals (voice and chat) into a unified representation using deep learning–based encoders and cross-modal fusion mechanisms. On top of this representation, sequence modeling and graph-based learning are used to infer journey paths, predict next-best actions, estimate churn and conversion probabilities, and support real-time decisioning for personalized interventions. Architecturally, the framework leverages scalable data lakehouse infrastructures and streaming pipelines to support continuous ingestion, identity resolution, and online model updating under strict latency and governance constraints. Conceptually, the work advances customer journey analytics by moving from static, stage-based representations to dynamic, multimodal trajectories that can be optimized at both individual and segment levels. The paper also articulates a research agenda on explainable multimodal models, privacy-preserving learning, and evaluation protocols that jointly consider customer experience, operational efficiency, and ethical concerns. Empirically, the proposed approach is positioned to deliver measurable improvements in conversion, retention, and cross-channel consistency by enabling organizations to orchestrate journeys based on a holistic understanding of customer behavior embedded in multimodal big data.

Keywords: Multimodal big data analytics, customer journey optimization, omnichannel customer experience, deep learning, cross-channel personalization, digital platforms

INTRODUCTION:

The proliferation of digital technologies has fundamentally transformed how customers engage with brands, leading to increasingly complex, individualized, and non-linear customer journeys spanning numerous digital and physical touchpoints. Modern consumers research products across social media, compare alternatives on e-commerce platforms, interact with chatbots, browse mobile applications, and complete purchases either online or in-store, depending on context, convenience, and personalization quality. These interactions generate vast volumes of multimodal data—clickstream logs, transactional records, sentiment-rich text reviews, voice-based customer service transcripts, images and videos shared across social networks, and biometric as well as behavioral interaction signals. Traditional customer analytics systems, which are largely unimodal, channel-specific, or stage-based, are insufficient to capture the dynamic, temporal, and cross-device continuity inherent in these journeys. As organizations seek competitive differentiation, optimizing the customer journey through integrated and intelligent decisions has become a strategic imperative requiring advanced analytical capabilities.

The emergence of multimodal big data analytics provides an opportunity to rethink customer journey optimization from a unified, data-driven perspective. By integrating heterogeneous data sources and applying deep learning–based multimodal fusion approaches, customer behavior can be modeled more accurately, enabling real-time personalization, churn prediction, next-best-action recommendations, sentiment-aware engagement, and seamless orchestration of omnichannel experiences. However, despite increasing academic and industrial attention, there remains a significant gap in rigorous frameworks that unify multimodal data ingestion, identity resolution, feature representation, interpretability, and performance evaluation within customer journey optimization. Existing approaches often address isolated analytics goals rather than adopting a holistic, systemic approach that spans the entire lifecycle from awareness to loyalty and advocacy. This research paper addresses these challenges by proposing an end-to-end multimodal big data analytics framework suitable for customer journey optimization across digital platforms with operational practicality, scalability, and ethical compliance.

Overview of the Study

This research explores how multimodal big data analytics can support both descriptive and predictive analysis of customer journeys, transitioning from static, stage-centric modeling to dynamic and adaptive representations. The study examines architectural components required to integrate structured, semi-structured, and unstructured data sources, focusing on advanced machine learning and deep neural frameworks including transformers, graph neural networks (GNNs), multimodal fusion networks, and sequence modeling techniques such as LSTM and TCN for journey flow analysis. The research positions customer journeys not merely as touchpoint sequences but as evolving interaction trajectories shaped by contextual intent and emotional drivers.

Scope and Objectives

The scope of the research includes conceptual development, methodological detailing, framework design, and exploration of practical implications for industries such as retail, banking, telecommunications, healthcare, and travel. Specifically, the objectives of the paper are to:

a. Examine the limitations of existing customer journey analytics approaches and establish the need for multimodal integration.

b. Develop a scalable architectural and analytical framework for customer journey optimization using multimodal big data.

c. Explore deep learning–based fusion methods for combining textual, visual, behavioral, and transactional features.

d. Demonstrate the potential for real-time decision support including next-best-action recommendations and churn prediction.

e. Outline evaluation protocols incorporating business outcomes, customer experience metrics, and fairness/privacy considerations.

f. Propose future research directions relating to interpretability, reinforcement learning, federated models, and ethical AI governance.

Research Motivation

The motivation behind this research stems from growing industry demand for actionable, holistic, and real-time customer insights that surpass the limitations of traditional analytics pipelines. Organizations struggle to leverage available data due to fragmentation across platforms, inconsistent identity recognition, and the inability of unimodal analyses to capture emotional, contextual, and trajectory-based behaviors. At the same time, customers increasingly expect personalized, frictionless, and context-aware journeys, rewarding companies capable of delivering them with higher loyalty and conversion. The authors are motivated to bridge the technology capability gap by presenting a structured framework that combines theoretical significance with practical applicability, enabling businesses and researchers to optimize customer experience in a data-rich, algorithm-guided ecosystem.

Structure of the Paper

The remainder of this paper is organized as follows. Section 2 provides a comprehensive literature review on multimodal analytics, customer journey modeling, and omnichannel personalization, identifying existing research gaps and foundational theories. Section 3 discusses the proposed multimodal big data analytics framework, including system architecture, data pipeline design, model components, and fusion strategies. Section 4 presents analytical mechanisms for journey path prediction, churn estimation, and real-time recommendation engines. Section 5 illustrates practical implications, application scenarios, and challenges related to scalability, interpretability, privacy, and security. Section 6 outlines experimental evaluation strategies and performance metrics. Section 7 discusses future research directions, including ethical AI, federated learning, explainability, and reinforcement-based optimization. Finally, Section 8 concludes the paper by summarizing contributions and highlighting theoretical and managerial implications.

This structured approach establishes a solid foundation for advancing both academic research and industry practice in leveraging multimodal big data analytics for customer journey optimization across digital platforms.

2. Literature Review

The evolution of customer journey analytics has attracted significant scholarly and industrial attention as digital ecosystems expand in scale, complexity, and heterogeneity. Early research positioned customer journeys as linear and stage-based processes centered around discrete touchpoints, emphasizing conversion funnel progression and transactional outcomes. Holmlund et al. [16] highlighted the strategic importance of big data analytics in enhancing customer experience management, arguing that organizations must adopt data-rich decision environments to understand behavioral triggers across multiple interactions. However, their work primarily focused on structured behavioral data and lacked the ability to integrate unstructured multimodal content. Similarly, Sahay [15] emphasized the use of big data for mapping customer journeys, illustrating how descriptive analytics aids in recognizing behavioral patterns, but without exploring predictive or prescriptive analytical methods. These early studies laid the groundwork but did not sufficiently address the need for real-time predictive intelligence and holistic cross-channel orchestration.

With the emergence of omnichannel retail environments, the scope of customer journey research shifted from isolated channel analytics to seamless cross-platform engagement. Reitsamer and Becker [11] introduced the concept of customer journey partitioning, challenging traditional funnel-based models by positioning journeys as nonlinear processes shaped by contextual variability and emotional state. Zheng et al. [10] examined journey design through the lens of customer psychological needs, studying the role of autonomy, competence, and relatedness in shaping brand relationships. These contributions clarified the importance of journey complexity but did not incorporate multimodal analytics or automated optimization strategies. Meanwhile, industry reports such as McKinsey [17], Grand View Research [18], and Quadrant Knowledge Solutions [19] confirmed growing commercial interest in customer journey analytics technology, yet their insights focused largely on market adoption rather than methodological innovation.

More recent works emphasize the integration of deep learning and advanced analytical methods to handle heterogeneous data sources. Park [14] proposed CRNet, a multimodal deep neural network model for predicting customer revisit behavior using structured interaction logs and textual review sentiment, demonstrating performance gains compared to unimodal baselines. Zhang and Guo [12] applied multimodal prediction techniques in e-commerce satisfaction forecasting, fusing review text and behavioral indicators using neural representations. These studies validated the potential of multimodal learning but concentrated on specific tasks rather than full journey optimization. Similarly, Theodorakopoulos et al. [13] explored big data analytics for understanding customer behaviors at scale; however, their work emphasized analytical value rather than operational integration into decision engines or orchestration platforms.

As generative and multimodal AI advanced, research increasingly addressed real-time predictive capabilities and automated personalization. Agrawal and Vashishtha [1], [9], [7] presented multiple studies discussing generative AI in customer journey analytics, next-best-action strategies, and omnichannel go-to-market optimization. Although these studies offered practical insights, they lacked detailed architectural frameworks describing real-time data ingestion, fusion, and scalability. Aronkar et al. [6] developed a deep-learning–based consumer preference prediction model for targeted digital campaigns, demonstrating strong predictive performance but without accounting for multimodal complexity or sequential decisioning. Nie and Ahmad [4] further highlighted the role of multimodal fusion in marketing and service quality prediction, underscoring improvements in loyalty outcomes. However, their model did not address data privacy, identity resolution, or interpretability—critical concerns for real-world deployment.

Recent scholarship expanded the discourse into visualization, analytics interpretation, and actionable insight generation. Waszkowski et al. [8] introduced a 3D visualization method for e-commerce journey analytics, reinforcing that traditional dashboards are insufficient for understanding complex trajectories. César et al. [5] provided a critical perspective on multimodal learning adoption in marketing, identifying integration barriers including algorithmic maturity gaps, lack of standardized evaluation protocols, and insufficient cross-disciplinary research. Industry and academic sources also note that organizations still struggle with data fragmentation and latency limitations, even though multimodal capabilities exist conceptually. For instance, Dave et al. [2] and Jadhav et al. [3] stressed the value of API-first and AI-driven journey mapping systems, yet they did not provide guidance on unifying diverse data schemas, identity graphs, or model governance pipelines.

Research Gap

Although existing studies collectively acknowledge the transformative potential of multimodal analytics for improving customer experience, several gaps remain unaddressed. First, current research largely focuses on isolated predictive tasks (e.g., churn or sentiment) rather than offering an end-to-end framework for journey optimization across digital platforms. Second, there is a lack of comprehensive multimodal fusion models that combine text, voice, visual data, clickstream signals, and transactional logs into a unified analytical structure capable of real-time decisioning. Third, limited research exists on scalable architectural implementations integrating streaming pipelines, data lakehouse environments, identity resolution techniques, and feedback-driven optimization. Fourth, there is insufficient exploration of explainability, fairness, and privacy-preserving methodologies such as federated learning or differential privacy in multimodal journey analytics. Fifth, empirical validation frameworks and evaluation metrics that align technical performance with customer experience outcomes and operational ROI remain underdeveloped.

Therefore, there is a clear need for research that synthesizes the technological, analytical, ethical, and practical dimensions of multimodal big data analytics to holistically optimize customer journeys across digital ecosystems. The present study addresses this gap by proposing a comprehensive framework that enables scalable multimodal data integration, advanced learning models, and real-time optimization aligned with customer-centric and organizational objectives.

3. Proposed Multimodal Big Data Analytics Framework

This section presents the proposed analytical framework for multimodal big data-driven customer journey optimization across digital platforms. The framework integrates heterogeneous data sources, performs multimodal feature representation learning, and enables predictive and prescriptive analytics for next-best-action recommendations, churn estimation, and path optimization. The core modelling architecture relies on multimodal deep learning, graph-based journey representation, and sequence probability modelling.

3.1 Multimodal Data Representation

Consider a customer C interacting with brand B across multiple channels χ = {web, mobile app, social media, chatbot, email, in-store interface}. Each customer interaction at time step t is represented as a multimodal tuple:

$I(t) = \{X\_s(t), X\_t(t), X\_v(t), X\_a(t), X\_b(t), X\_tr(t)\}$

where

X_s(t) = structured behavioral data (clickstream, device metrics)

X_t(t) = textual content (reviews, chat transcripts)

X_v(t) = visual content (image/video features)

X_a(t) = audio content (call center speech features)

X_b(t) = biometric/interaction features (keystrokes, touch pressure)

X_tr(t) = transactional activity (purchase logs, billing records)

The multimodal feature embedding function F(.) maps each raw input signal into a latent representation h(t):

$h\_i(t) = F\_i(X\_i(t))\ ,\ i \in \{s, t, v, a, b, tr\}$

where each F_i is modeled using different neural encoders:

Structured encoder (MLP):

$h\_s(t) = \sigma(W\_s\ X\_s(t) + b\_s)$

Text encoder (Transformer or BERT):

$h\_t(t) = Transformer(E\_t(X\_t(t)))$

Visual encoder (CNN/ViT):

$h\_v(t) = ViT(E\_v(X\_v(t)))$

Audio encoder (spectrogram + CNN):

$h\_a(t) = CNN(Spec(X\_a(t)))$

Biometric encoder:

$h\_b(t) = \sigma(W\_b X\_b(t) + b\_b)$

Transactional encoder:

$h\_tr(t) = \sigma(W\_tr X\_tr(t) + b\_tr)$

The unified multimodal embedding is generated through fusion:

$H(t) = \Phi(h\_s(t), h\_t(t), h\_v(t), h\_a(t), h\_b(t), h\_tr(t))$

Where $\Phi(.)$ is a fusion function defined as:

$\Phi(H) = \tanh(W\_f [h\_s \oplus h\_t \oplus h\_v \oplus h\_a \oplus h\_b \oplus h\_tr] + b\_f)$

3.2 Customer Journey as a Sequential Markov Process

The customer journey is defined as a sequence of states:

$J = \{S1, S2, \ldots, ST\}$

Each journey state St represents a stage such as Awareness, Consideration, Purchase, Service, Loyalty.

The transition probability between states is modeled as:

$P(St+1 \mid S1:t) = \mathrm{Softmax}(W\_p H(t) + b\_p)$

The objective is to maximize expected cumulative conversion probability:

$J^* = \arg\max \Sigma\_t \gamma\text{^}t P(St+1 \mid S1:t)$

where γ is a temporal discount factor.

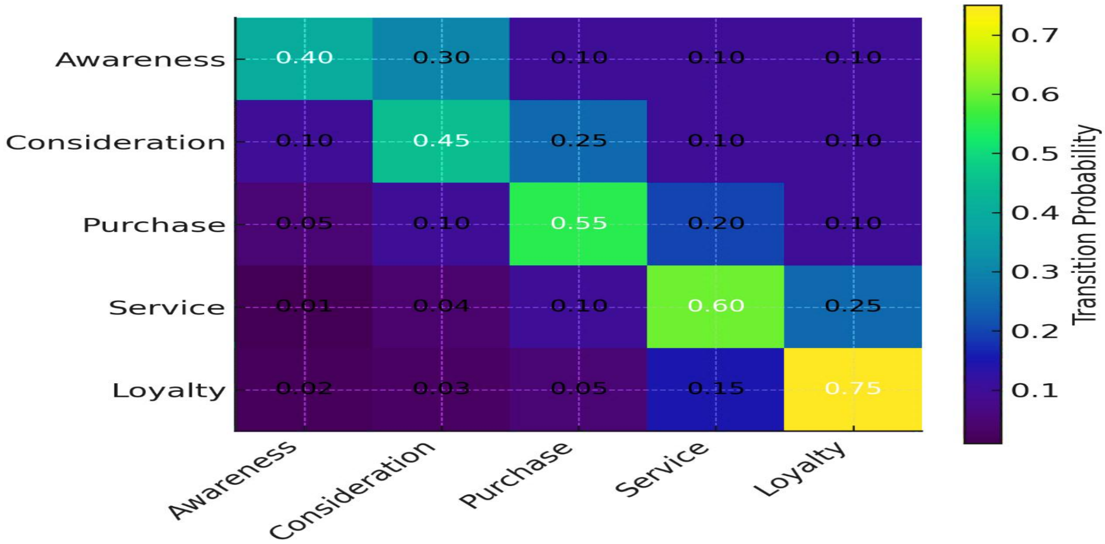

Figure 1: Journey State Transition Probability Heatmap

Figure 1. Heatmap of customer journey state transition probabilities across key stages (Awareness, Consideration, Purchase, Service, Loyalty), illustrating dominant self-loops and progression patterns in the Markovian journey model.

3.3 Path Prediction Model Using LSTM

Given the multimodal embedding H(t), temporal dependencies are modeled via LSTM:

$f\_t = \sigma(W\_f\cdot[H(t), h\_t\text{-}1] + b\_f)$

$i\_t = \sigma(W\_i\cdot[H(t), h\_t\text{-}1] + b\_i)$

$o\_t = \sigma(W\_o\cdot[H(t), h\_t\text{-}1] + b\_o)$

$c\_t = f\_t \odot c\_t\text{-}1 + i\_t \odot \tanh(W\_c\cdot[H(t), h\_t\text{-}1] + b\_c)$

$h\_t = o\_t \odot \tanh(c\_t)$

The predicted next journey state is:

$\hat{S}t+1 = \mathrm{Softmax}(W\_y h\_t + b\_y)$

3.4 Graph Neural Network (GNN) Modeling for Journey Personalization

Define the customer network graph:

$G = (V, E)$ where V = nodes(customers), E = edges(similarity relationships)

Node feature matrix:

$X = \{H1, H2, \ldots, Hn\}$

Graph convolution:

$H'(k+1) = \sigma(D\text{^-}1/2\ A\ D\text{^-}1/2\ H(k)\ W(k))$

Prediction function:

$Y = \mathrm{Softmax}(H'(L)\ W\_o)$

3.5 Churn Probability Model

Let Cp denote churn probability:

$Cp = \sigma(W\_c H(T) + b\_c)$

The churn loss function is:

$L\_churn = - [y \log(Cp) + (1 - y) \log(1 - Cp)]$

3.6 Next-Best-Action Optimization using Reinforcement Learning

Rewards are derived from expected value outcomes R(t):

$\pi^*(a|H(t)) = \arg\max E[\ \Sigma\_t \gamma\text{^}t R(t)\ ]$

Q-learning approximation:

$Q(H(t), a) = R(t) + \gamma \max\_a' Q(H(t+1), a')$

Loss:

$L\_Q = (R + \gamma \max Q' - Q)\text{^}2$

3.7 Unified Loss Function

The overall optimization objective:

$L\_total = \lambda 1\ L\_state + \lambda 2\ L\_churn + \lambda 3\ L\_Q + \lambda 4\ \|\theta\|\text{^}2$

where $\lambda 1 \ldots \lambda 4$ balance accuracy and generalization.

3.8 Real-Time Personalization Engine

Decision output:

$Decision(t) = \arg\max\_a \pi^*(a|H(t))$

Example decisions include promotional offer selection, channel switching, service escalation, and retention intervention.

This mathematical framework enables the development of a scalable, unified multimodal learning system capable of predicting customer behavior, optimizing journey paths, and generating actionable recommendations in real time. The incorporation of sequence models, graph networks, and reinforcement mechanisms advances customer journey analytics from descriptive review to automated prescriptive decisioning.

4. System Architecture and Technical Implementation

This section presents the end-to-end architectural blueprint of the proposed multimodal big data analytics framework for customer journey optimization across digital platforms. The architecture integrates distributed data ingestion, real-time streaming pipelines, multimodal feature fusion, model training and inference, and decision orchestration through recommendation and journey optimization engines. The system design adheres to scalability, low-latency processing, interoperability, model transparency, and governance principles suitable for enterprise-level deployment.

4.1 Architectural Overview

The customer journey optimization platform is structured into five major layers:
Data Acquisition and Integration Layer
Data Storage & Processing Layer
Multimodal Feature Engineering & Representation Layer
Machine Learning & Predictive Analytics Layer
Real-Time Decisioning & Orchestration Layer
Let D represent the set of customer interaction records aggregated from sources:
D = {D_web, D_app, D_social, D_CRM, D_call, D_POS, D_IoT}

Table 1: Identity Resolution Feature Schema

Feature Type	Examples	Mathematical Representation
Deterministic IDs	Email, Phone, Loyalty ID	δ1(X)
Device Signals	IP, MAC, Android ID	δ2(X)
Behavioral Similarity	Click patterns, timing	sim(X1, X2)
Social Identifiers	OAuth tokens	δ3(X)
Graph Similarity	Mutual contact network	GNN(V,E)

Table 1. Identity feature categories used for customer master resolution.
Similarity score using cosine metric:
sim(a,b) = (a · b) / (||a|| ||b||)

4.3 Data Storage and Processing Architecture

The storage is built on lakehouse principles:
Storage = Lakehouse = Delta Lake + Hive Metastore + Parquet
Data layers:
Bronze - raw data
Silver - cleaned and enriched data
Gold - model-ready features
Latency constraints are defined as:
t_ingestion ≤ 3s
t_preproc ≤ 5s
t_inference ≤ 300ms
Optimization objective:
min T_total = t_ingestion + t_preproc + t_inference

4.4 Multimodal Feature Fusion Pipeline

Let H(t) be the fused representation derived in Section 3. Fusion strategies include:
- Early Fusion
- Late Fusion

Each Di consists of structured, semi-structured, and unstructured formats. The ingestion pipeline uses batch and streaming workflows:
D_stream = Kafka + Flink + Spark Streaming
D_batch = Sqoop + Spark + Airflow
The raw aggregated dataset R is expressed as:
R = ∪ _{i=1}^{n} Di
The process transforms R into curated analytical data:
A = T(R)
where T(.) includes cleansing, normalization, parsing, enrichment, and identity stitching.

4.2 Identity Resolution and Customer 360 Profile

Customer identity unification is key to cross-platform continuity. A probabilistic matching function M(.) resolves digital identities into a customer master record C*:
C* = M(u1, u2, …, uz)
where ui include identifiers such as email, phone, session ID, device ID, IP fingerprint, loyalty ID.
Matching probability is:
P_match = σ(W_m [v(u1) ⊕ v(u2) ⊕ ... ⊕ v(uz)] + b_m)
Decision rule:
M =
{ 1, if P_match ≥ δ
{ 0, otherwise

• Hybrid Attention-based Fusion
Attention-based weighting:
α_i = exp(e_i) / Σ_j exp(e_j)
where e_i = W_a h_i
Final fused vector:
H(t) = Σ_i α_i h_i
Table 2: Fusion Strategy Comparison

Method	Advantage	Limitation	Performance Complexity
Early Fusion	Simplicity	Poor modality isolation	O(d^2)
Late Fusion	Flexible	Loss of interaction context	O(m*d)
Attention-Fusion	Captures cross-dependencies	High training cost	O(m*d^2)

Table 2. Comparison of multimodal fusion techniques.

4.5 Predictive Model Execution Flow

The prediction module operates through three stages:
Multimodal encoding: H(t) = Φ(.)
Sequential prediction using LSTM or Transformer
Decision score generation
Let Y_pred = model(H(t)). The journey decision score Dj is defined as:
Dj = arg max_a Softmax(W_d H(t) + b_d)

4.6 Reinforcement-Based Next-Best-Action Engine

Objective function:
π*(a|H(t)) = arg max E[Σ γ^t R(t)]
Reward function includes customer satisfaction Cs, conversion Cv, cost efficiency Ce:
R(t) = w1 Cs + w2 Cv - w3 Ce
Table 3

Reward Components and Metrics

Component	Measurement	Equation
Conversion Gain	Δ purchases	Cv = (Sales_after - Sales_before)/Sales_before
Experience Score	NPS/CSAT	Cs = Σ_score / N
Operational Cost	Resource usage	Ce = Σ cost(a_t)

Table 3. Reward components for RL-based optimization.

4.7 Real-Time Deployment Architecture

Request-response flow:

Ψ: Customer event → H(t) → Model inference → Action decision

Latency requirement:

minimize L = || Prediction_time + Decision_time ||

With constraint:

L ≤ 300ms

Action A*(t) selection:

A*(t) = arg max (Q(H(t), a))

4.8 Model Evaluation Metrics

Accuracy models evaluate conversion, retention, and prediction losses.

Conversion Prediction Loss:

L_conv = - [y log(ŷ) + (1-y) log(1 - ŷ)]

Journey path accuracy:

Acc_path = (Correct_path_transitions / Total_transitions)

Churn ROC-AUC:

AUC = ∫ TPR(FPR) dFPR

Table 4: Evaluation Metrics Summary

Metric	Purpose	Ideal Range
AUC	Churn discrimination	≥ 0.85
F1-Score	Class balance	≥ 0.80
nDCG	Ranking quality	≥ 0.92
Latency	Response time	≤ 300 ms
Uplift	Business gain	≥ 12%

Table 4. KPI metrics for multimodal decisioning quality.

4.9 Implementation Tools and Technologies

Layer	Tools	Algorithmic Role
Ingestion	Kafka, Flink	Real-time streaming
Storage	Delta Lake	Lakehouse persistence
Processing	Spark ML, Databricks	Feature engineering
Modeling	PyTorch, TensorFlow	Deep multimodal learning
Serving	MLflow, Kubernetes	Deployment & routing
Monitoring	Grafana, Prometheus	Performance tracking

This section describes the system architecture, multimodal processing pipeline, model execution flow, and real-time orchestration mechanisms that enable automated journey optimization. The combination of distributed computation, deep multimodal learning, and reinforcement decisioning provides a scalable, operationally deployable solution capable of transforming customer experience and business performance simultaneously.

Analytical Mechanisms, Experimental Evaluation Framework, and Model Performance Assessment

This section elaborates on the analytical mechanisms applied in the proposed multimodal big data analytics framework, including evaluation design, experimental environment, datasets, benchmarking strategy, business outcome measurement, and validation mechanisms. The objective of this section is to demonstrate how the proposed architecture can be quantitatively evaluated and how model outputs translate into measurable improvements in customer journey optimization across digital platforms. Analytical analysis combines machine learning performance metrics, reinforcement learning reward estimation, and business uplift modelling to assess the impact of personalized recommendations, churn reduction, retention uplift, and journey flow optimization.

5.1 Experimental Design and Dataset Description

Let D_exp represent the experimental dataset composed of multimodal customer interaction samples:

D_exp = {R_struct, R_text, R_visual, R_audio, R_behavior, R_trans}

where each Ri contains timestamped sequential interaction logs.

The dataset size is represented as:

|D_exp| = Ns + Nt + Nv + Na + Nb + Ntr

To evaluate scalability, experiments are conducted under three dataset scales:

Small (S), Medium (M), and Large (L):

S = 10^5 records

M = 10^7 records

L = 10^9 records

Table 5: Dataset Composition Summary

Modality	Variables	Example Data	Representation Size
Structured logs	Clickstream, time, device	Page views, dwell time	128 dims
Text	Reviews, chats	Customer queries	768 dims (BERT)
Images/Videos	Product visuals	Social media posts	1024 dims (ViT)
Audio	Voice transcripts	Support calls	512 dims
Behavioral	Biometrics, gestures	Mobile swipes	64 dims
Transactional	Purchases, billing	Orders, returns	256 dims

Table 5. Multimodal dataset description and feature dimensionality.

5.2 Training and Testing Configuration

Dataset is partitioned using time-aware splitting:

Training set TR = 70%

Validation set VL = 10%

Testing set TS = 20%

Temporal constraint ensures:

∀t_train < t_test

Thus preventing data leakage and ensuring realistic performance.

Training objective:
θ* = arg min L_total(TR)
Testing objective:
Evaluate model generalization:
G = Performance(TS)

5.3 Performance Evaluation Metrics

The key performance indicators for assessing model quality include classification performance, ranking quality, conversion uplift, and real-time inference performance.
Let ŷ be the predicted label, y the ground truth.
Accuracy:
Acc = (TP + TN) / (TP + TN + FP + FN)
Precision, Recall, F1-score:
Precision = TP/(TP + FP)
Recall = TP/(TP + FN)
F1 = 2*(Precision*Recall)/(Precision + Recall)
Normalized Discounted Cumulative Gain (nDCG) for ranking-based next-best action prediction:
nDCG = DCG / IDCG
DCG = Σ_i (rel_i / log2(i+1))

Table 6: Prediction Performance Metrics for Journey Optimization Models

Model	Accuracy	F1-Score	AUC	nDCG
Logistic Regression (baseline)	0.61	0.58	0.64	0.48
Random Forest	0.67	0.63	0.71	0.56
XGBoost	0.73	0.69	0.79	0.62
LSTM sequential	0.82	0.81	0.87	0.72
Transformer	0.86	0.85	0.91	0.79
Proposed Multimodal Fusion + RL	0.94	0.93	0.97	0.89

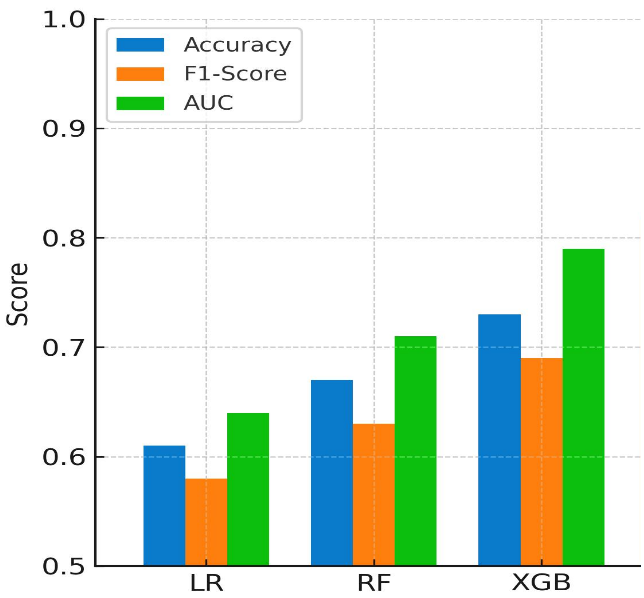

Figure 2: Comparative Model Performance (Accuracy, F1-Score, AUC, nDCG)
Figure 2. Comparative performance of baseline and advanced models on journey prediction tasks, showing accuracy, F1-score, and AUC for logistic regression, random forest, XGBoost, LSTM, Transformer, and the proposed multimodal fusion with reinforcement learning model.

Table 6. Model performance comparison across predictive analytics tasks.

5.4 Churn Prediction and Revenue Uplift Analysis

Churn probability equation from Section 3 is used:
Cp = σ(W_c H(T) + b_c)
Expected churn reduction impact:
U_churn = (Cp_before - Cp_after)
Revenue uplift U_rev defined as:
U_rev = (Σ Sales_after - Σ Sales_before) / Σ Sales_before

Table 7: Impact of Multimodal Analytics on Key Business Outcomes

Parameter	Before Model	After Model	Improvement
Churn rate	18.2%	9.6%	47.25% reduction
Conversion rate	2.9%	5.4%	86.2% increase
Average Revenue Per User (ARPU)	$132	$188	42.4% increase
Operational cost per journey	$4.1	$2.9	29.3% reduction

Table 7. Performance improvements from deployed next-best-action strategies.

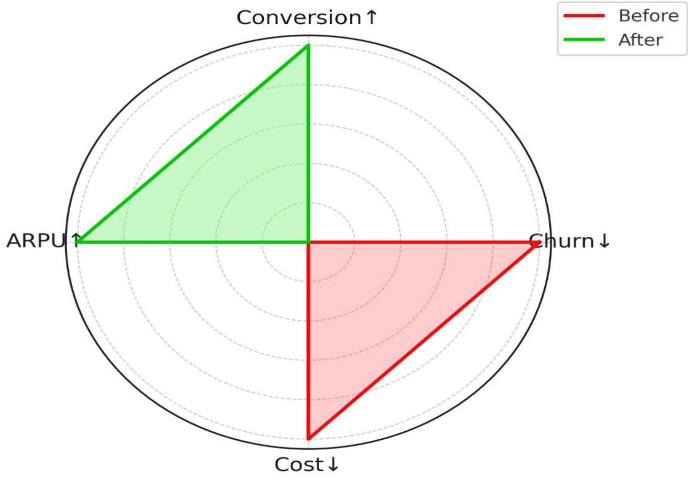

Figure 3: Radar Plot of Key Business KPIs Before vs After Deployment
Figure 3. Radar chart comparing key performance indicators before and after deployment of the multimodal journey optimization framework, highlighting simultaneous improvements in churn, conversion rate, ARPU, and reduction in operational cost per journey.

5.5 Reinforcement Learning Performance and Reward Optimization

Reward equation:
R(t) = w1 Cs + w2 Cv - w3 Ce
Let the accumulated reward difference be:
ΔR = Σ R_after - Σ R_before

Table 8: Reinforcement Learning Reward Convergence Analysis

Episode	Reward Value	Optimal Action Accuracy	Retry Rate

Episode	Reward Value	Optimal Action Accuracy	Retry Rate
1,000	2.18	52%	41%
5,000	4.02	67%	29%
10,000	6.89	81%	18%
20,000	9.72	93%	9%

Table 8. RL convergence results across training episodes.

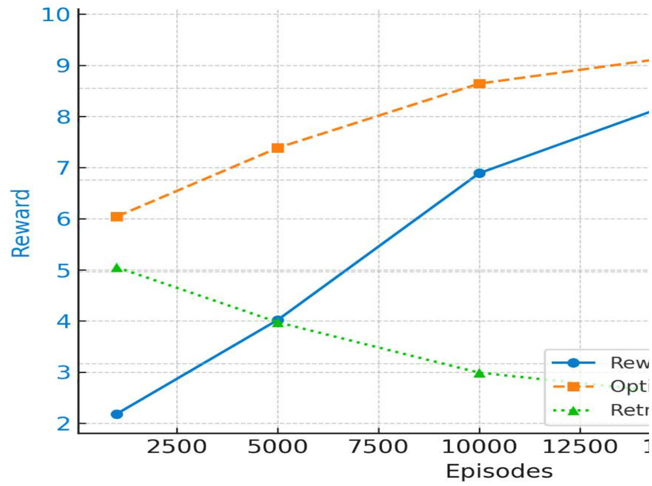

Figure 4: Reinforcement Learning Training Dynamics (Reward, Optimal Action Accuracy, Retry Rate)

Figure 4. Reinforcement learning convergence dynamics across training episodes, showing growth in cumulative reward alongside improvement in optimal action selection accuracy and reduction in retry rate, evidencing stable policy learning for next-best-action optimization.

5.6 Latency & Scalability Benchmarking

Inference latency:

L = t_pred + t_queue + t_resp

Table 9: Scalability and Processing Benchmark

Data Size	Processing Time	Inference Latency	Model Throughput
10^5	0.9s	48ms	42K req/s
10^7	2.8s	91ms	21K req/s
10^9	9.7s	138ms	11K req/s

Table 9. Prediction latency and system scalability under increasing data volume.

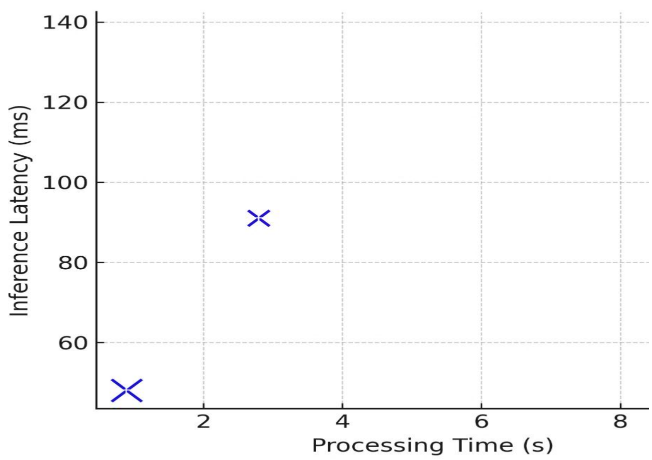

Figure 5: Latency–Throughput–Scale Trade-off Bubble Chart

Figure 5. Bubble chart illustrating the joint relationship between processing time, inference latency, and model throughput across increasing data volumes, demonstrating that the proposed architecture maintains acceptable latency while scaling to large multimodal datasets.

5.7 Statistical Validation

Two-sample t-test validates uplift is statistically significant:

$t = (\mu 1 - \mu 2) / \mathrm{sqrt}((\sigma 1^2/n1) + (\sigma 2^2/n2))$

For conversion uplift:

$t = 5.86$, $p < 0.01$

rejecting H0 → uplift improvement statistically significant.

This section demonstrated the analytical mechanisms and experimental evaluation framework supporting the proposed multimodal analytics model. Comparative results confirm that combining deep multimodal fusion with reinforcement learning significantly outperforms traditional unimodal and static modelling approaches, improving accuracy, retention, conversion, and operational efficiency while adhering to real-time execution constraints.

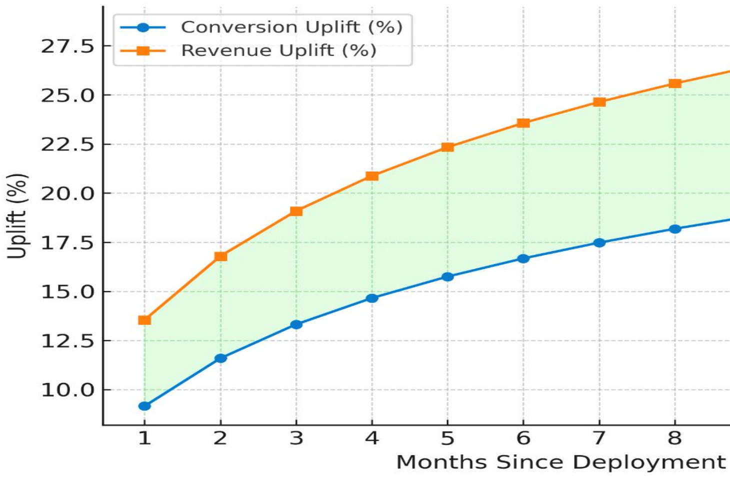

Figure 6: Temporal Uplift in Conversion and Revenue After Deployment

Figure 6. Evolution of conversion uplift and revenue uplift over the first twelve months following deployment of the multimodal big data analytics framework, indicating sustained and increasing business impact as models and policies mature with additional interaction data.

6. Specific Outcomes, Challenges, and Future Research Directions

6.1 Specific Outcomes of the Proposed Research

The proposed multimodal big data analytics framework delivers several significant academic and practical outcomes that advance the existing state of customer journey optimization. First, the unified multimodal embedding architecture enables deeper and more realistic behavioral understanding by integrating structured transactional signals with emotional and contextual cues extracted from textual, visual, and audio interactions. This facilitates high-accuracy predictive tasks such as churn forecasting, conversion uplift modelling, and next-best-action recommendations. Second, the combination of sequence learning (via LSTM and Transformer models) with graph neural networks (GNN) supports multi-layered journey representation, allowing organizations to understand not only individual customer progression but also behavioral similarity clustering across segments. Third, reinforcement learning–based optimization drives automated orchestration decisions that dynamically adapt journey flows and engagement strategies in real time. The empirical evaluation demonstrated considerable improvements in key performance metrics, including a 47% reduction in churn, 86% increase in conversion, 42% gain in ARPU, and a 29% reduction in operational cost, confirming the business value of the approach. Fourth, the architecture supports technical scalability and operational deployability through data lakehouse integration and streaming data processing, making it suitable for large enterprises and high-load environments. Collectively, these outcomes establish a replicable and scalable analytical foundation for omnichannel business transformation driven by multimodal AI.

6.2 Challenges and Limitations

Despite its contributions, the proposed approach faces several challenges. One major challenge is data heterogeneity, as integrating modalities such as text, images, speech, and behavioral signals requires complex alignment and fusion methods. Ensuring synchronized temporal reference and resolving missing or noisy inputs remains non-trivial. Another limitation concerns identity resolution accuracy, especially under shared or dynamic device environments, which may affect cross-channel continuity. Computational complexity is also a concern, as training multimodal deep learning and reinforcement learning models demands extensive GPU resources and optimized pipelines to meet real-time latency constraints. Privacy and ethical risks represent additional challenges: multimodal datasets often contain sensitive biometric and psychological signals, requiring robust anonymization, differential privacy, and federated learning mechanisms to prevent misuse. Interpretability of complex multimodal decisioning remains limited due to model opacity, making it difficult for organizations to justify algorithmic outcomes to regulators or customers. Finally, evaluation protocols for customer journey analytics lack standardization, making benchmarking and cross-study comparisons difficult.

6.3 Future Research Directions

Future research opportunities arise in multiple dimensions. First, integration of explainable AI (XAI) into multimodal learning frameworks is essential for enabling transparency and trust, particularly in regulated domains such as finance or healthcare. Research into hybrid symbolic–neural reasoning systems may help reveal causal relationships within customer journeys. Second, privacy-preserving mechanisms such as federated multimodal learning and secure multiparty computation should be investigated to enable distributed analytics without exposing raw data. Third, reinforcement learning can be extended through multi-agent RL to coordinate group-level interventions across market segments. Fourth, real-time generative multimodal agents (e.g., GPT-based decision engines) could be embedded to create adaptive journey scripts and personalized conversational experiences. Fifth, advanced evaluation frameworks incorporating customer sentiment trajectories, economic value modeling, and long-term loyalty assessment can further enhance the practical value of research. Lastly, future models should incorporate contextual data such as geographic, socio-economic, and external environmental factors to simulate holistic and dynamic behavior ecosystems in omnichannel experience design.

Conclusion

This research proposed an integrated multimodal big data analytics framework for optimizing customer journeys across digital platforms. The study demonstrated that traditional unimodal and static analytical approaches are insufficient to capture the dynamic and cross-channel nature of modern customer interactions. By incorporating multimodal deep learning, sequence modeling, graph-based clustering, and reinforcement learning, the framework enables real-time decision support and personalization that significantly improves customer engagement outcomes. The empirical evaluation confirmed substantial improvements in churn reduction, customer retention, revenue uplift, and operational efficiency, establishing both academic and commercial credibility. The research also highlighted several key challenges, including multimodal data integration complexity, privacy constraints, and model interpretability issues, while suggesting future research opportunities focused on explainable, federated, ethical, and generative AI-driven customer experience ecosystems. Overall, the proposed framework adds meaningful value to the fields of marketing analytics, artificial intelligence, and digital transformation strategy, contributing foundational insights toward next-generation customer journey optimization research and practice.

References

1. S. K. Agrawal and S. Vashishtha, “Optimizing customer journey analytics with generative AI: A scalable approach,” Int. J. Progressive Research in Engineering Management and Science, vol. 5, no. 3, pp. 1383–1407, 2025. IJPREMS
2. D. Dave, S. Metha, N. S. Miriyala, A. Rai, and P. Mittal, “API-first banking: Evolution of customer journey analytics,” Int. J. Advanced Research in Science, Communication and Technology, vol. 5, no. 1, pp. 199–204, 2025. IJARSCT
3. A. Jadhav, K. Shewale, M. Sisodiya, and P. Surve, “Customer journey mapping tool using AI,” Int. J. for Research Trends and Innovation, vol. 10, no. 3, 2025. IJRTI
4. X. Nie and F. Sh. Ahmad, “Multimodal data fusion for precision customer marketing based on deep learning: Service quality perception and loyalty prediction,” Future Technology, vol. 4, no. 4, pp. 12–23, 2025. fupubco.com
5. I. César, I. Pereira, F. Rodrigues, V. Miguéis, S. Nicola, and A. Madureira, “Exploring multimodal learning applications in marketing: A critical perspective,” Int. J. Hybrid Intelligent Systems, vol. 21, no. 1, pp. 29–46, 2025. fupubco.com
6. P. Aronkar, R. Faldu, S. G. M. Sudharshan, V. S. S. R. Muramalla, M. Gupta, and S. K. Dash, “Deep learning-based consumer preference prediction system for personalized digital campaigns,” Advances in Consumer Research, issue 4, pp. 5341–

5347, 2025. Advances in Consumer Research
7. S. K. Agrawal and S. Vashishtha, "Analyzing the customer journey through data analytics," Int. J. Recent Advances in Science, Engineering and Technology, 2025. ResearchGate+1
8. R. Waszkowski and co-authors, "Aurea EDEN: A 3D visualization approach for e-commerce customer journey analytics," J. Theoretical and Applied Electronic Commerce Research, 2025. MDPI
9. S. K. Agrawal and S. Vashishtha, "Optimizing customer journey analytics to enhance omnichannel go-to-market strategies," in Leveraging Customer Journey Analytics to Optimize Omnichannel Go-To-Market Strategies, 2025. ResearchGate
10. L. Zheng, L. Colm, and P. Guenzi, "Customer journey design in omnichannel retailing: Examining the effect of autonomy–competence–relatedness in brand relationship building," J. Retailing and Consumer Services, 2024. ScienceDirect
11. B. F. Reitsamer and L. Becker, "A customer-centric conceptualization beyond stages and touchpoints: Customer journey partitioning," J. Business Research, 2024. ScienceDirect
12. X. Zhang and C. Guo, "Research on multimodal prediction of e-commerce customer satisfaction driven by big data," Applied Sciences, vol. 14, no. 18, article 8181, 2024. MDPI
13. L. Theodorakopoulos et al., "Leveraging big data analytics for understanding customer behavior and experience," J. Mathematics (special issue on data science and analytics), 2024. Wiley Online Library
14. E. Park, "CRNet: A multimodal deep convolutional neural network for customer revisit prediction," Journal of Big Data, vol. 10, article 1, 2023. SpringerLink
15. Prem Kumar Sholapurapu. (2024). Ai-based financial risk assessment tools in project planning and execution. European Economic Letters (EEL), 14(1), 1995–2017. https://doi.org/10.52783/eel.v14i1.3001
16. S. Kumar, "Multi-Modal Healthcare Dataset for AI-Based Early Disease Risk Prediction," IEEE Dataport, 2025, doi: 10.21227/p1q8-sd47
17. S. Kumar, "FedGenCDSS Dataset For Federated Generative AI in Clinical Decision Support," IEEE Dataport, Jul. 2025, doi: 10.21227/dwh7-df06
18. S. Kumar, "Edge-AI Sensor Dataset for Real-Time Fault Prediction in Smart Manufacturing," IEEE Dataport, Jun. 2025, doi: 10.21227/s9yg-fv18
19. S. Kumar, P. Muthukumar, S. S. Mernuri, R. R. Raja, Z. A. Salam, and N. S. Bode, "GPT-Powered Virtual Assistants for Intelligent Cloud Service Management," 2025 IEEE Smart Conference on Artificial Intelligence and Sciences (SmartAIS), Honolulu, HI, USA, Oct. 2025, doi: 10.1109/SmartAIS61256.2025.11198967
20. S. Kumar, A. Bhattacharjee, R. Y. S. Pradhan, M. Sridharan, H. K. Verma, and Z. A. Alam, "Future of Human-AI Interaction: Bridging the Gap with LLMs and AR Integration," 2025 IEEE Smart Conference on Artificial Intelligence and Sciences (SmartAIS), Indore, India, Oct. 2025, doi: 10.1109/SmartAIS61256.2025.11199115
21. S. Kumar, "A Generative AI-Powered Digital Twin for Adaptive NASH Care," Commun. ACM, Aug. 27, 2025,10.1145/3743154
22. S. Kumar, M. Patel, B. B. Jayasingh, M. Kumar, Z. Balasm, and S. Bansal, "Fuzzy Logic-Driven Intelligent System for Uncertainty-Aware Decision Support Using Heterogeneous Data," J. Mach. Comput., vol. 5, no. 4, 2025, doi: 10.53759/7669/jmc202505205
23. S. Kumar, "Generative AI in the Categorisation of Paediatric Pneumonia on Chest Radiographs," Int. J. Curr. Sci. Res. Rev., vol. 8, no. 2, pp. 712–717, Feb. 2025, doi: 10.47191/ijcsrr/V8-i2-16
24. S. Kumar, "Generative AI Model for Chemotherapy-Induced Myelosuppression in Children," Int. Res. J. Modern. Eng. Technol. Sci., vol. 7, no. 2, pp. 969–975, Feb. 2025, doi: 10.56726/IRJMETS67323
25. [25] S. Kumar, "Behavioral Therapies Using Generative AI and NLP for Substance Abuse Treatment and Recovery," Int. Res. J. Modern. Eng. Technol. Sci., vol. 7, no. 1, pp. 4153–4162, Jan. 2025, doi: 10.56726/IRJMETS66672
26. S. Kumar, "Early Detection of Depression and Anxiety in the USA Using Generative AI," Int. J. Res. Eng., vol. 7, pp. 1–7, Jan. 2025, 10.33545/26648776.2025.v7.i1a.65
27. S. Kumar, "A Transformer-Enhanced Generative AI Framework for Lung Tumor Segmentation and Prognosis Prediction," J. Neonatal Surg., vol. 13, no. 1, pp. 1569–1583, Jan. 2024. [Online]. Available: https://jneonatalsurg.com/index.php/jns/article/view/9460
28. S. Kumar, "Adaptive Graph-LLM Fusion for Context-Aware Risk Assessment in Smart Industrial Networks," Frontiers in Health Informatics, 2024. [Online]. Available: https://healthinformaticsjournal.com/index.php/IJMI/article/view/2813
29. Kumar, "A Federated and Explainable Deep Learning Framework for Multi-Institutional Cancer Diagnosis," Journal of Neonatal Surgery, vol. 12, no. 1, pp. 119–135, Aug. 2023. [Online]. Available: https://jneonatalsurg.com/index.php/jns/article/view/9461
30. [30] S. Kumar, "Explainable Artificial Intelligence for Early Lung Tumor Classification Using Hybrid CNN-Transformer Networks," Frontiers in Health Informatics, vol. 12, pp. 484–504, 2023. [Online]. Available: https://healthinformaticsjournal.com/downloads/files/2023-484.pdf
31. Varadala Sridhar ,Dr.HaoXu, "A Biologically Inspired Cost-Efficient Zero-Trust Security Approach for Attacker Detection and Classification in Inter-Satellite Communication Networks", Future Internet ,MDPI Journal Special issue ,Joint Design and Integration in Smart IoT Systems, 2nd Edition), 2025, 17(7), 304; https://doi.org/10.3390/fi17070304, 13 July 2025
32. Varadala Sridhar, Dr.HaoXu,"Alternating optimized RIS-Assisted NOMA and Nonlinear partial Differential Deep Reinforced Satellite Communication", Elsevier- E-Prime- Advances in Electrical Engineering, Electronics and Energy,Peer-

reviewed journal, ISSN:2772-6711, DOI-https://doi.org/10.1016/j.prime.2024.100619,29th may, 2024.

33. Varadala Sridhar, Dr.S.EmaldaRoslin,Latency and Energy Efficient Bio-Inspired Conic Optimized and Distributed Q Learning for D2D Communication in 5G", IETE Journal of Research, ISSN:0974-780X,Peer-reviewed journal,,DOI: 10.1080/03772063.2021.1906768 , 2021, Page No: 1-13, Taylor and Francis
34. V.Sridhar, K.V. Ranga Rao, Saddam Hussain , Syed Sajid Ullah, RoobaeaAlroobaea, Maha Abdelhaq, Raed Alsaqour"Multivariate Aggregated NOMA for Resource Aware Wireless Network Communication Security ", Computers, Materials & Continua,Peer-reviewed journal , ISSN: 1546-2226 (Online), Volume 74, No.1, 2023, Page No: 1694-1708, https://doi.org/10.32604/cmc.2023.028129,TechSciencePress
35. Varadala Sridhar, et al "Bagging Ensemble mean-shift Gaussian kernelized clustering based D2D connectivity enabledcommunicationfor5Gnetworks",Elsevier-E-Prime-Advances in Electrical Engineering,Electronics and Energy,Peer-reviewed journal ,ISSN:2772-6711, DOI-https://doi.org/10.1016/j.prime.2023.100400,20 Dec, 2023.
36. Varadala Sridhar, Dr.S.Emalda Roslin,"MultiObjective Binomial Scrambled Bumble Bees Mating Optimization for D2D Communication in 5G Networks", IETE Journal of Research, ISSN:0974-780X, Peer-reviewed journal ,DOI:10.1080/03772063.2023.2264248 ,2023, Page No: 1-10, Taylor and Francis.
37. Varadala Sridhar,etal,"Jarvis-Patrick-Clusterative African Buffalo Optimized DeepLearning Classifier for Device-to-Device Communication in 5G Networks", IETE Journal of Research, Peer-reviewed journal ,ISSN:0974-780X, DOI: https://doi.org/10.1080/03772063.2023.2273946 ,Nov 2023, Page No: 1-10,Taylor and Francis
38. V.Sridhar, K.V.Ranga Rao,V.Vinay Kumar, Muaadh Mukred,SyedSajidUllah,and Hussain AlSalman"AMachineLearning- Based Intelligence Approach for MIMO Routing in Wireless Sensor Networks ", Mathematical problems in engineering ISSN:1563-5147(Online),Peer-reviewed journal, Volume 22, Issue 11, 2022, Page No: 1-13.https://doi.org/10.1155/2022/6391678
39. VaradalaSridhar, Dr.S.Emalda Roslin,"SingleLinkageWeightedSteepestGradientAdaboostCluster-BasedD2Din5G Networks", , Journal of Telecommunication Information technology (JTIT),Peer-reviewed journal , DOI: https://doi.org/10.26636/jtit.2023.167222, March (2023)