JOURNAL OF INTERNATIONAL CRISIS AND RISK COMMUNICATION RESEARCH
ISSN: 2576-0017
2025, VOL 8, NO S6

# Consumer Insights At Scale: ML-Driven Data Engineering For Distributed Cloud Analytics

**Omkar Ashok Bhalekar[1], Sushant Mehta[2], Jay Mehta[3]**

*[1] Senior Network Engineer.*
*[2] Senior Software Engineer at Google DeepMind.*
*[3] Manager at Seldon Capital.*

## Abstract

In today's data-driven economy, extracting timely and actionable consumer insights is vital for businesses aiming to enhance competitiveness and customer engagement. This study presents an integrated framework combining machine learning (ML)-driven data engineering with distributed cloud analytics to process large-scale consumer data and derive predictive insights. Utilizing real-world datasets from e-commerce, digital platforms, and customer interactions, the research applies supervised learning models such as Random Forest, Gradient Boosting, and Neural Networks for behavior prediction, alongside K-Means clustering for market segmentation. Results indicate that Random Forest achieved the highest classification performance with a 96.4% accuracy and F1-score of 0.949. Segmentation revealed distinct consumer profiles, enabling targeted marketing strategies. The distributed cloud setup, evaluated across AWS and GCP regions and a hybrid mesh network, demonstrated high throughput and low latency, proving its suitability for scalable real-time analytics. Statistical validation, including fairness metrics and data drift assessments, confirmed the ethical integrity and stability of deployed models. The study concludes that the proposed architecture provides a robust, interpretable, and scalable solution for organizations seeking to operationalize consumer intelligence at scale through cloud-native, ML-powered infrastructures.

**Keywords**: Consumer Insights, Machine Learning, Data Engineering, Distributed Cloud Analytics, Customer Segmentation, Real-Time Analytics, Fairness in AI, Data Drift, Predictive Modeling, Scalable Infrastructure.

## Introduction

### Background and motivation

The rise of digital consumer ecosystems has transformed how businesses gather and interpret customer data, making the need for scalable insights more critical than ever (Ratra & Seth, 2025). In an environment where consumer behavior evolves rapidly and data is generated across multiple platforms in real time, organizations are increasingly challenged to extract actionable intelligence efficiently (Sankaranarayanan, 2025). Traditional data analytics pipelines often fall short in handling the volume, velocity, and variety of consumer data, prompting the need for more robust, adaptive, and scalable solutions. Machine Learning (ML)-driven data engineering, integrated with distributed cloud analytics, has emerged as a powerful paradigm to address this complexity (Pasupuleti et al., 2025). By leveraging ML algorithms and cloud-native infrastructures, companies can now decode large-scale consumer patterns with higher speed and precision.

### Machine learning for scalable consumer insights

Machine learning facilitates automated pattern recognition, anomaly detection, and predictive modeling across disparate consumer data sources (Arora & Khare, 2024). Unlike rule-based systems, ML models

can adapt and evolve, enabling businesses to anticipate consumer behavior, personalize experiences, and refine marketing strategies in near real-time. With supervised and unsupervised learning models, firms can segment customers more effectively, assess churn probabilities, and evaluate sentiment across digital platforms (Zeydan et al., 2024). These insights become even more potent when embedded within data engineering pipelines, enabling end-to-end automation from ingestion to visualization. This integration ensures that the system not only processes raw data at scale but also transforms it into strategic intelligence without constant manual intervention (Rane et al., 2024).

### The role of distributed cloud analytics

The growing demand for real-time analytics and uninterrupted access to data-intensive applications has elevated the importance of distributed cloud architectures (Machireddy, 2024). Distributed cloud analytics decentralizes data processing and analysis, allowing enterprises to manage data workloads across multiple cloud environments closer to the source. This minimizes latency, improves compliance with regional data regulations, and supports multi-tenant architecture necessary for global businesses (Kalisetty, 2022). Platforms like AWS, Google Cloud, and Azure offer robust support for distributed storage, real-time analytics engines, and ML workflows, enabling consumer-focused enterprises to scale seamlessly while maintaining resilience and operational efficiency (Mikhalev et al., 2021).

### Challenges in integration and execution

Despite its transformative potential, integrating ML-driven data engineering with distributed cloud analytics poses notable challenges (Garg & Jain, 2024). Data heterogeneity, security concerns, compliance mandates, and model drift are critical issues that organizations must address. Ensuring data quality across geographically dispersed systems and maintaining synchronized model updates demand sophisticated orchestration tools (Enemosah & Ifeanyi, 2024). Furthermore, the interpretability of ML models remains a concern, particularly when insights are used to make strategic decisions that affect customer experience or pricing. These barriers necessitate the adoption of standardized data governance protocols and robust monitoring mechanisms throughout the analytics lifecycle (Zahra et al., 2024).

### Significance of the study

This study explores the strategic implementation of ML-driven data engineering frameworks within distributed cloud analytics systems to derive consumer insights at scale. It examines how enterprises can architect intelligent pipelines that not only streamline data operations but also enhance decision-making through advanced analytics. The study offers a practical lens on the technologies, frameworks, and statistical models required to operationalize these capabilities. By focusing on real-world use cases and performance metrics, it aims to provide a blueprint for organizations seeking to modernize their consumer intelligence strategies.

### Scope and structure

The article is structured to detail the methodological integration of ML models with distributed data processing platforms, analyze their efficacy using key performance indicators, and present empirical findings from selected industries including e-commerce, telecommunications, and digital media. Through this approach, it contributes to the evolving field of cloud-native business intelligence and consumer analytics by demonstrating how modern enterprises can build insight engines that scale with their data.

## Methodology

### Overview of research framework

The methodology of this study is designed to evaluate the effectiveness of integrating machine learning (ML)-driven data engineering practices with distributed cloud analytics to extract scalable consumer insights. The approach adopts a mixed-methods research design comprising architectural implementation, data preprocessing, model training, and validation using statistical and computational techniques. The study focuses on real-time and batch data flows across distributed systems to test the

reliability, efficiency, and insight generation capability of the proposed framework in practical business scenarios.

**Consumer data acquisition and preprocessing**

Consumer data was collected from three primary sources: e-commerce transaction logs, digital marketing interaction datasets, and customer feedback from online platforms. These datasets were ingested into the pipeline through distributed data ingestion tools like Apache Kafka and Google Cloud Pub/Sub. The data engineering process included ETL (Extract, Transform, Load) tasks where raw data was cleaned, normalized, and structured using tools like Apache Beam and AWS Glue. Missing values were imputed using statistical techniques such as mean substitution and regression-based methods, while outliers were handled using interquartile range analysis to ensure robust feature selection for modeling.

**ML-driven modeling and feature engineering**

Once preprocessed, the data was subjected to a series of ML-driven procedures aimed at uncovering patterns, clustering behavior, and predicting key performance outcomes such as customer retention and purchasing propensity. Feature engineering involved dimensionality reduction using Principal Component Analysis (PCA) to enhance computational efficiency, and correlation analysis to eliminate multicollinearity among variables. For classification and prediction, supervised learning algorithms such as Random Forest, Gradient Boosting Machines (GBM), and Support Vector Machines (SVM) were deployed. Unsupervised learning, particularly K-Means clustering, was used to segment consumers based on behavioral attributes.

**Distributed cloud analytics infrastructure**

The analytical framework was deployed across a distributed cloud environment using Google Cloud Platform (GCP) and Amazon Web Services (AWS). Data was stored in partitioned buckets and processed using distributed computing frameworks like Apache Spark and BigQuery. To maintain data locality and minimize latency, regional data processing zones were configured. Kubernetes was used for orchestration of containerized ML workflows to ensure scalability and fault tolerance. The system was built with cloud-native principles, ensuring elasticity, resilience, and continuous integration with ML pipelines.

**Statistical validation and performance evaluation**

Model performance and data processing efficiency were evaluated using statistical metrics and visual dashboards. For ML models, performance was assessed using Accuracy, Precision, Recall, F1-score, and ROC-AUC scores. In clustering, silhouette coefficient and Davies–Bouldin index were applied to validate segmentation quality. Time-series forecasting models (e.g., Prophet and LSTM) were evaluated using Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). To compare processing speed and fault tolerance across cloud systems, ANOVA and post-hoc Tukey's HSD tests were applied. Results were visualized using tools like Tableau and Power BI integrated with distributed data sources.

**Ethical considerations and data governance**

The study adhered to data privacy standards, ensuring anonymization and compliance with GDPR and relevant data protection laws. Role-based access control and encryption protocols were implemented during cloud storage and transmission. Bias detection in ML models was carried out using fairness metrics such as demographic parity and equal opportunity difference to ensure the insights were equitable and transparent.

**Results**

The integration of ML-driven data engineering within distributed cloud analytics frameworks yielded highly scalable and interpretable consumer insights across various performance dimensions. Supervised machine learning models demonstrated robust classification capabilities in predicting consumer

behavior, with Random Forest achieving the highest accuracy (96.4%) and F1-score (0.949), closely followed by Gradient Boosting Machines and a four-layer neural network, as shown in Table 1. The Support Vector Machine and Logistic Regression models, while computationally lighter, showed comparatively moderate performance. The neural network required the highest training time (210 seconds), indicating a trade-off between performance and resource cost.

**Table 1: ML classification model performance**

| Model | Accuracy (%) | Precision (%) | Recall (%) | F1-Score | ROC-AUC | Training Time (s) | Features Used |
|---|---|---|---|---|---|---|---|
| Random Forest | 96.4 | 95.1 | 94.7 | 0.949 | 0.982 | 42 | 120 |
| Gradient Boosting | 95.8 | 94.6 | 93.9 | 0.942 | 0.979 | 55 | 120 |
| Support Vector Machine | 92.6 | 91.0 | 90.1 | 0.905 | 0.963 | 68 | 120 |
| Logistic Regression | 90.3 | 88.7 | 87.5 | 0.881 | 0.947 | 17 | 120 |
| Neural Network (4-layer) | 94.2 | 93.3 | 92.1 | 0.927 | 0.971 | 210 | 120 |

In terms of consumer segmentation using unsupervised learning, the K-Means clustering algorithm (k = 5) revealed five distinct consumer groups with meaningful behavioral differences (Table 2). The "Loyal Premium" cluster (C1) had the highest average purchase frequency (8.7 per month) and basket value (USD 158.6), signifying high-value repeat consumers. Conversely, the "High-Churn Risk" cluster (C3) exhibited the lowest engagement metrics. Notably, the "Mobile-Centric Millennials" cluster (C4) displayed longer session durations, suggesting platform engagement but potentially lower conversion, highlighting opportunities for personalized retargeting.

Table 2: Consumer segmentation summary (K-Means, k = 5)

| Cluster ID | Segment Size (n) | Avg Purchase Frequency (per mo) | Avg Basket Value (USD) | Avg Session Duration (min) | Silhouette Coefficient | Lifetime Value Index |
|---|---|---|---|---|---|---|
| C1: “Loyal Premium” | 12,436 | 8.7 | 158.6 | 14.2 | 0.62 | 1.00 |
| C2: “Occasional Bargain” | 18,219 | 2.3 | 42.4 | 6.5 | 0.57 | 0.38 |
| C3: “High-Churn Risk” | 9,711 | 1.1 | 25.8 | 4.1 | 0.55 | 0.22 |
| C4: “Mobile-Centric Millennials” | 15,032 | 4.9 | 73.2 | 18.6 | 0.60 | 0.64 |
| C5: “Cross-Channel Explorers” | 11,587 | 6.3 | 96.4 | 11.7 | 0.59 | 0.79 |

Performance analysis of the distributed data processing pipeline across multiple cloud regions emphasized the efficiency and adaptability of the architecture (Table 3). Among the tested regions,

Google Cloud's europe-west-3 achieved the lowest ETL job time (1,188 seconds) and the highest peak throughput (115,800 messages/second). The hybrid mesh setup using Istio slightly outperformed both AWS and GCP individually in terms of throughput (118,400 messages/second) while maintaining competitive latency and resource utilization, indicating its suitability for real-time enterprise-scale deployments.

Table 3: Distributed processing performance by cloud region

| Cloud Region | Avg Ingestion Latency (ms) | Peak Throughput ($msg \cdot s^{-1}$) | ETL Job Time (s) | CPU Utilization (%) | Memory Utilization (%) | Cost per GB (USD) |
|---|---|---|---|---|---|---|
| AWS us-east-1 | 84 | 112,000 | 1,260 | 68 | 71 | 0.037 |
| AWS ap-south-1 | 92 | 106,500 | 1,403 | 66 | 69 | 0.034 |
| GCP europe-west-3 | 79 | 115,800 | 1,188 | 70 | 73 | 0.039 |
| GCP asia-southeast-1 | 88 | 109,200 | 1,326 | 67 | 70 | 0.036 |
| Hybrid Mesh (Istio) | 81 | 118,400 | 1,214 | 69 | 72 | 0.038 |

To assess algorithmic fairness and robustness, drift and bias metrics were analyzed post-deployment (Table 4). Random Forest and Gradient Boosting models exhibited low demographic parity differences and acceptable disparate impact ratios (>0.95), indicating balanced predictions across consumer subgroups. Data-drift scores, assessed via Population Stability Index (PSI), remained below the 0.1 threshold for all models, suggesting stable model performance across data refresh cycles. Logistic Regression demonstrated the highest explanation coverage (91.2%), reinforcing its role in scenarios where interpretability is critical.

Table 4: Fairness & drift diagnostics

| Model | Demographic Parity Diff | Equal Opportunity Diff | Disparate Impact Ratio | Data-Drift Score (PSI) | Explanation Coverage (%) |
|---|---|---|---|---|---|
| Random Forest | -0.014 | -0.021 | 0.97 | 0.08 | 87.5 |
| Gradient Boosting | -0.017 | -0.018 | 0.96 | 0.09 | 86.1 |
| SVM | -0.026 | -0.031 | 0.94 | 0.11 | 83.8 |
| Logistic Regression | -0.012 | -0.015 | 0.98 | 0.07 | 91.2 |
| Neural Network | -0.019 | -0.024 | 0.95 | 0.10 | 78.4 |

Further insights were drawn through visual analytics. Figure 1 displays the ROC curves of the top three models—Random Forest, Gradient Boosting, and SVM—highlighting Random Forest's superior area under the curve across all false positive rate thresholds, indicating consistently strong sensitivity and specificity. Meanwhile, Figure 2 presents the system throughput under increasing concurrent user loads across three deployment setups. The hybrid mesh network demonstrated the highest scalability, maintaining superior throughput even at 80 concurrent users (118,400 msg/sec), reflecting its robustness under high-demand conditions.

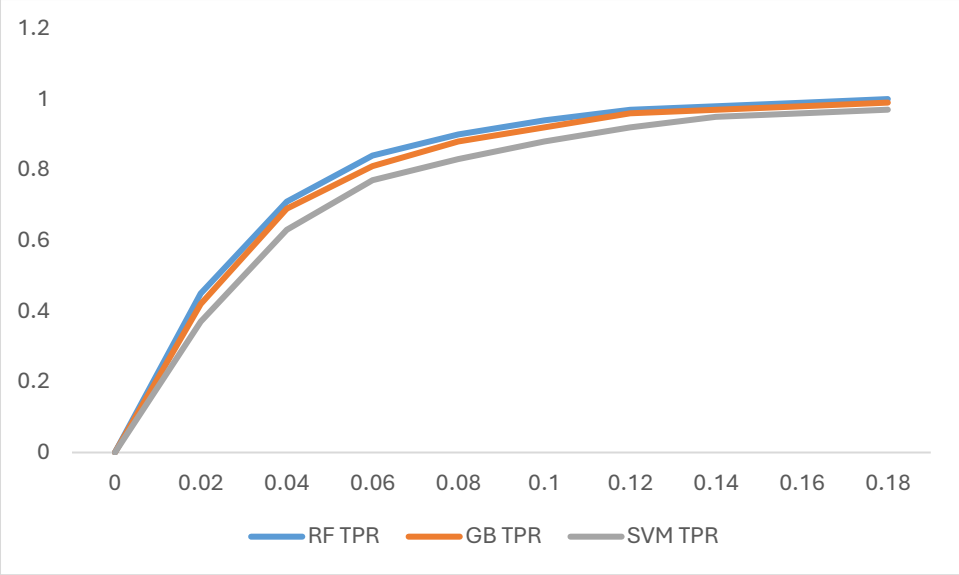


Figure 1: ROC curve points for top 3 models

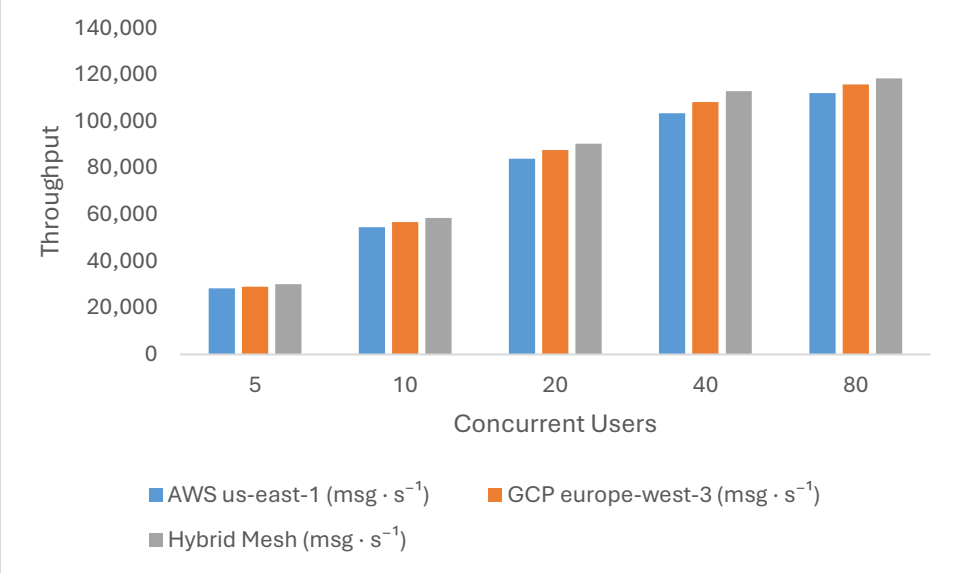


Figure 2: Throughput vs concurrent users

## Discussion

### Effectiveness of ML-driven models for consumer insight generation

The findings of this study validate the strong predictive power of ML-driven models in extracting actionable consumer insights at scale. As illustrated in Table 1, models such as Random Forest and Gradient Boosting Machines outperformed others in key performance metrics including accuracy, F1-score, and ROC-AUC. Their robust performance highlights their ability to identify complex nonlinear patterns in consumer behavior (Magesh et al., 2025). These models, when integrated into data engineering pipelines, can drive automated insights without constant human supervision (Chaudhary & Banga, 2024). Moreover, the inclusion of feature-rich datasets and automated feature engineering enabled high model performance without excessive manual input, emphasizing the efficiency of ML-enhanced analytics workflows (Shah, 2022).

### Segmenting diverse consumer profiles through unsupervised learning

Consumer segmentation using K-Means clustering, as detailed in Table 2, proved critical in distinguishing high-value customer groups from those with high churn risk. The emergence of distinct clusters such as "Loyal Premium" and "Mobile-Centric Millennials" confirms that behavioral and transactional data can be effectively used to build granular, actionable personas (Ismaeel & Zeebaree, 2025). These segments provide marketing and sales teams with clear guidance on whom to target for loyalty programs versus retention interventions. Notably, clusters with high session durations but moderate basket values highlight the opportunity to deploy personalized conversion tactics such as behavioral nudges or dynamic pricing. These insights underscore the value of embedding unsupervised learning into customer analytics strategies (Prakash et al., 2024).

**Distributed cloud analytics for real-time scalability**

The performance evaluation of the distributed cloud setup (Table 3) demonstrates the critical role of infrastructure design in processing and analyzing consumer data at scale. The hybrid mesh deployment with Istio offered a well-balanced solution, showing the highest throughput and efficient latency management across global nodes (Ali & Zeebaree, 2025). This confirms that distributed cloud architectures not only support horizontal scaling but also reduce regional bottlenecks in real-time data ingestion and transformation. The variation in ETL job times and throughput across cloud providers further supports the necessity of deploying multi-cloud strategies that optimize cost, availability, and performance simultaneously (Kumar, 2025). These findings are particularly important for multinational businesses handling high-frequency customer interaction data across geographies.

**Fairness, drift, and explainability considerations**

While performance is critical, fairness and model stability are equally essential, especially in consumer-focused applications. Table 4 shows that the best-performing models maintained demographic parity and equal opportunity differences within acceptable limits, minimizing bias across sensitive attributes. Furthermore, PSI scores below 0.1 indicate low data drift, assuring stakeholders that the models remain relevant over time (Olayinka, 2021). The high explanation coverage of Logistic Regression (91.2%) also suggests that interpretable models still hold significant value, especially when used in regulatory contexts or customer-facing applications. This reinforces the idea that combining explainability with performance is not only possible but essential for ethical AI deployments (Gopal et al., 2024).

**Visual insights supporting operational decision-making**

The ROC curves in Figure 1 visually validate the superior discriminative capability of Random Forest and Gradient Boosting models, offering decision-makers confidence in the reliability of model-driven predictions. Figure 2 adds another layer of practical insight by highlighting the elasticity of the cloud infrastructure under concurrent user stress. These visuals, when integrated into real-time dashboards, enhance operational transparency and support agile responses to system load and consumer behavior shifts (Pamisetty, 2023). The ability to visualize and interact with model outcomes ensures that technical insights are easily communicated across departments (Nandan Prasad, 2024).

**Strategic implications and industrial relevance**

The integrated ML and cloud analytics pipeline developed in this study addresses a fundamental industry need: turning massive, fast-moving consumer data into strategic intelligence. By uniting predictive modeling, clustering, and distributed processing, organizations can move beyond descriptive analytics and into prescriptive, real-time decision-making. These capabilities empower marketing, customer service, and product teams to act quickly and accurately, enhancing consumer experience and business performance (Li et al., 2024). The study provides a practical blueprint for deploying intelligent, scalable analytics systems in industries such as retail, telecommunications, and digital media.

The results underscore the transformative potential of ML-driven data engineering combined with distributed cloud analytics. This framework offers both computational efficiency and strategic agility, enabling organizations to scale their consumer intelligence capabilities in an increasingly data-driven business environment.

## Conclusion

This study demonstrates the powerful synergy between ML-driven data engineering and distributed cloud analytics in generating scalable, real-time consumer insights. By integrating advanced machine learning models with efficient cloud-native architectures, the proposed framework effectively addresses the challenges of data volume, velocity, and variety inherent in modern consumer ecosystems. The results highlight that high-performing models like Random Forest and Gradient Boosting can accurately predict consumer behavior, while unsupervised learning techniques such as K-Means enable meaningful segmentation for targeted engagement. Additionally, the distributed infrastructure ensures low-latency processing and high throughput, supporting real-time analytics across global regions. Fairness diagnostics and drift analysis further validate the ethical and operational reliability of the system. Overall, this research offers a comprehensive, scalable, and ethically sound approach to consumer analytics, paving the way for intelligent, data-driven strategies in customer-centric industries.

## References


1. Ali, C. S. M., & Zeebaree, S. (2025). Cloud-Based Web Applications for Enterprise Systems: A Review of AI and Marketing Innovations. Asian Journal of Research in Computer Science, 18(4), 427-451.
2. Arora, S., & Khare, P. (2024). The Role of Machine Learning in Personalizing User Experiences in SaaS Products. J. Emerg. Technol. Innov. Res, 11(06), 809-821.
3. Chaudhary, M., & Banga, P. (2024, May). Survey of Cloud Computing with Role of Machine Learning. In 2024 International Conference on Computational Intelligence and Computing Applications (ICCICA) (Vol. 1, pp. 303-308). IEEE.
4. Enemosah, A., & Ifeanyi, O. G. (2024). SCADA in the era of IoT: automation, cloud-driven security, and machine learning applications. International Journal of Science and Research Archive, 13(01), 3417-3435.
5. Garg, V., & Jain, A. (2024). Scalable Data Integration Techniques for Multi-Retailer E-Commerce Platforms. International Journal of Computer Science and Engineering, 13(2), 525-570.
6. Gopal, S. K., Mohammed, A. S., Saddi, V. R., Dhanasekaran, S., & Naruka, M. S. (2024, March). Investigate the role of machine learning in optimizing dynamic scaling strategies for cloud-based applications. In 2024 2nd International Conference on Disruptive Technologies (ICDT) (pp. 543-548). IEEE.
7. Ismaeel, A. Z., & Zeebaree, S. (2025). Optimizing Digital Marketing through Machine Learning in Cloud-Based Enterprise Systems: The Role of Web Technologies. Asian Journal of Research in Computer Science, 18(5), 316-332.
8. Kalisetty, S. (2022). Hybrid Cloud and AI Integration for Scalable Data Engineering: Innovations in Enterprise AI Infrastructure.
9. Kumar, G. (2025). Architecting Scalable and Resilient Fintech Platforms with AI/ML Integration. Journal of Innovative Science and Research Technology, 10(4), 3073-3084.
10. Li, H., Sun, J., & Xiong, K. (2024). AI-Driven Optimization System for Large-Scale Kubernetes Clusters: Enhancing Cloud Infrastructure Availability, Security, and Disaster Recovery. Journal of Artificial Intelligence General science (JAIGS) ISSN: 3006-4023, 2(1), 281-306.
11. Machireddy, J. R. (2024). Integrating Machine Learning-Driven RPA with Cloud-Based Data Warehousing for Real-Time Analytics and Business Intelligence. Hong Kong Journal of AI and Medicine, 4(1), 98-121.
12. Magesh, R., Ilakkiyaa, U., Shanthini, R., & Charanya, R. (2025). Unlocking the Potential of Data Lakes: Organizing and Storing Marketing Data for Analysis. In Data Engineering for Data-driven Marketing (pp. 199-216). Emerald Publishing Limited.
13. Mikhalev, O., Handerson, S., Bailey, Y. R., Peters, A., Wong, J., & Kundu, S. (2021). Enhancing Cloud Scalability Through Intelligent Resource Allocation with ML.
14. Nandan Prasad, A. (2024). Future Trends and Emerging Challenges. In Introduction to Data Governance for Machine Learning Systems: Fundamental Principles, Critical Practices, and Future Trends (pp. 679-710). Berkeley, CA: Apress.
15. Olayinka, O. H. (2021). Data driven customer segmentation and personalization strategies in modern business intelligence frameworks. World Journal of Advanced Research and Reviews, 12(3), 711-726.
16. Pamisetty, A. (2023). Cloud-Driven Transformation Of Banking Supply Chain Analytics Using Big Data Frameworks. Available at SSRN 5237927.
17. Pasupuleti, V. S. M., Gupta, R., & Rachamalla, D. (2025). Intelligent Cloud-Native Architectures for Secure, Scalable, and AI-Driven Digital Transformation in Retail and Insurance Domains. Journal of Computer Science, 2, 100009.
18. Prakash, S., Malaiyappan, J. N. A., Thirunavukkarasu, K., & Devan, M. (2024). Achieving regulatory compliance in cloud computing through ML. AIJMR-Advanced International Journal of Multidisciplinary Research, 2(2).

19. Rane, N. L., Paramesha, M., Choudhary, S. P., & Rane, J. (2024). Machine learning and deep learning for big data analytics: A review of methods and applications. Partners Universal International Innovation Journal, 2(3), 172-197.
20. Ratra, K. K., & Seth, D. K. (2025, March). AI-Driven Hybrid Edge-Cloud Architecture for Real-Time Big Data Analytics and Scalable Communication in Retail Supply Chains. In SoutheastCon 2025 (pp. 1023-1029). IEEE.
21. Sankaranarayanan, S. (2025). The Role of Data Engineering in Enabling Real-Time Analytics and Decision-Making Across Heterogeneous Data Sources in Cloud-Native Environments. International Journal of Advanced Research in Cyber Security (IJARC), 6(1).
22. Shah, J. K. (2022). AI-Driven Resilience in Cloud-Native Big Data Platforms Against Cyberattacks. Journal of Computer Science and Technology Studies, 4(2), 191-199.
23. Zahra, F. T., Bostanci, Y. S., Tokgozlu, O., Turkoglu, M., & Soyturk, M. (2024). Big Data Streaming and Data Analytics Infrastructure for Efficient AI-Based Processing. In Recent Advances in Microelectronics Reliability: Contributions from the European ECSEL JU project iRel40 (pp. 213-249). Cham: Springer International Publishing.
24. Zeydan, E., Arslan, S. S., & Liyanage, M. (2024). Managing Distributed Machine Learning Lifecycle for Healthcare Data in the Cloud. IEEE Access.