COLOMBO SCIENTIFIC PUBLISHING

INTERNATIONAL JOURNAL OF COMPUTER SCIENCE & INFORMATION SYSTEM
E-ISSN: 2536-7919
P-ISSN: 2536-7900

https://doi.org/10.55640/ijcsis/Volume11Issue08-07 PAGE NO:119-127

Edge-Cloud AI Computing: A Robust Framework for Real-Time Inference and Decision Automation

Dr. Arif Setiawan
Department of Artificial
Intelligence and Digital Technologies Indonesia

Dr. Maya Permata
Department of Machine
Intelligence and Computer Science Indonesia

ARTICLE INFO

Article history:

Submission: June,29 2026
Accepted: July 22, 2026
Published: August 20, 2026
VOLUME: Vol.11 Issue 08 2026

Keywords:
Edge-cloud computing, artificial intelligence, real-time inference, decision automation, federated learning, communication efficiency, model aggregation, distributed AI, reinforcement learning, resilient computing

ABSTRACT

The rapid deployment of artificial intelligence (AI) across distributed environments has created a need for computing architectures capable of simultaneously supporting low-latency inference, scalable model execution, communication efficiency, privacy, and reliable decision automation. Conventional cloud-centric AI architectures provide substantial computational capacity but can introduce communication delays, bandwidth dependency, privacy exposure, and service disruption risks when real-time decisions must be generated close to data sources. Edge-cloud AI computing addresses these limitations by distributing inference, coordination, and computational workloads across edge devices and cloud infrastructure. This research and review article develops a robust conceptual framework for real-time edge-cloud AI inference and decision automation by synthesizing research on federated learning, communication compression, heterogeneous model aggregation, blockchain-enabled healthcare systems, reinforcement learning, and distributed AI pipelines. The analysis identifies communication efficiency, heterogeneous computational capabilities, privacy preservation, adaptive resource allocation, and resilient inference coordination as the principal architectural requirements. The proposed framework integrates edge-level preprocessing and inference, adaptive communication, cloud-level model coordination, and automated decision orchestration. Findings indicate that compression and knowledge distillation can reduce communication burdens, while heterogeneous aggregation and adaptive learning mechanisms improve the practicality of distributed AI environments. Reinforcement learning further provides a mechanism for dynamic resource and decision optimization. However, the framework remains constrained by device heterogeneity, synchronization overhead, model inconsistency, security requirements, and the trade-off between inference accuracy and latency. The study establishes an integrated theoretical foundation for designing scalable and resilient edge-cloud AI systems capable of supporting real-time intelligent decision processes.

INTRODUCTION

The increasing deployment of AI-enabled applications has transformed inference from an isolated computational task into a distributed systems problem. Modern intelligent applications frequently generate data through geographically dispersed sensors, mobile devices, Internet-of-Things platforms, healthcare equipment, autonomous systems, and enterprise infrastructures. Processing all such data through centralized cloud platforms can provide substantial computational resources, but it can also create latency, bandwidth, privacy, and availability constraints. Edge-cloud AI computing therefore represents an architectural transition in which computational intelligence is distributed between resource-constrained edge environments and comparatively powerful cloud infrastructure.

The central challenge is not simply determining whether inference should occur at the edge or in the cloud. Rather, the challenge is determining which computation should occur where, when information should be exchanged, how models should be coordinated, and how decisions should remain reliable when communication or computing conditions change. This perspective is particularly important for real-time applications in which delayed inference may reduce the value of an otherwise accurate prediction. Kumar (2025) conceptualizes resilient edge-to-cloud inference pipelines as an architectural mechanism for supporting real-time decision systems, emphasizing the importance of resilience and coordinated inference across distributed computational layers.

Federated learning provides an important theoretical foundation for this architecture because it allows distributed participants to contribute to model development without requiring centralized collection of all raw data. Its application to healthcare informatics demonstrates the relevance of decentralized learning where data privacy and institutional boundaries are significant considerations (Xu et al., 2021). However, federated learning introduces its own communication and aggregation challenges. The resulting architecture must therefore address not only computation but also the efficiency and reliability of model exchange.

Communication efficiency is particularly important because edge devices may operate over bandwidth-limited or unreliable networks. Compression-based federated learning has consequently received significant attention, including autoencoder-based compression, rate adaptation, gradient compression, and sparsity-oriented mechanisms (Beitollahi and Lu, 2022; Cui et al., 2022; Mao, 2023; Sun et al., 2020). These approaches suggest that robust edge-cloud AI requires communication to be treated as an architectural resource rather than as a transparent supporting function.

This article develops a research-driven framework for edge-cloud AI computing focused on real-time inference and decision automation. Its objectives are to synthesize the provided literature, identify architectural requirements, formulate an integrated edge-cloud inference framework, examine its expected outcomes, and critically evaluate its limitations. The scope encompasses distributed AI inference, federated learning, communication optimization, heterogeneous model coordination, reinforcement learning, and privacy-oriented architectures. The study does not assume that all workloads should be moved to either edge or cloud; instead, it treats workload placement as a dynamic architectural decision.

LITERATURE REVIEW

2.1 Federated Learning as a Foundation for Distributed AI

Federated learning provides a decentralized learning paradigm in which participating devices or organizations retain local data while contributing model updates to a coordination process. Xu et al. (2021) demonstrate its relevance to healthcare informatics, where data may be distributed among institutions and subject to strong privacy constraints. This characteristic is directly applicable to edge-cloud architectures because edge devices can perform local learning or inference while the cloud coordinates broader model evolution.

Healthcare applications also reveal the importance of secure distributed data exchange. Salim and Park (2023) propose a federated learning-based approach for secure electronic health record sharing, demonstrating how federated mechanisms can be combined with security requirements in medical informatics. Lakhan (2023) further connects federated learning with blockchain-based privacy preservation and fraud management in healthcare IoMT environments. These studies collectively indicate that distributed AI cannot be evaluated purely through predictive accuracy; trust, privacy, data ownership, and system integrity must also be considered.

Myrzashova et al. (2023) provide a broader examination of blockchain and federated learning integration in healthcare, identifying challenges and opportunities associated with combining decentralized learning and distributed trust mechanisms. The implication for edge-cloud AI is that decentralized intelligence may require multiple control mechanisms: federated learning for model coordination, security mechanisms for protecting updates, and potentially distributed ledgers for accountability.

2.2 Communication-Efficient Distributed Intelligence

Communication overhead is a fundamental limitation of distributed AI. Beitollahi and Lu (2022) introduce FLAC, combining autoencoder compression with federated learning and convergence guarantees. This demonstrates that model communication can be compressed while preserving the mathematical conditions necessary for learning convergence.

Cui et al. (2022) examine optimal rate adaptation for federated learning under compressed communications. Their work is particularly relevant to edge-cloud systems because communication quality is not static. An architecture that assumes a constant transmission rate may perform poorly when edge connectivity changes. Adaptive communication therefore becomes an essential component of resilient inference pipelines.

Sun et al. (2020) address adaptive federated learning with gradient compression in uplink NOMA environments. Their results reinforce the importance of jointly considering wireless communication characteristics and federated optimization. Mao (2023), through the SAFARI framework, addresses sparsity-enabled federated learning under limited and unreliable communications. This perspective is especially relevant to edge environments where intermittent connectivity and constrained bandwidth may be normal rather than exceptional conditions.

Wu et al. (2022) approach communication efficiency from another direction by using knowledge distillation. Instead of transmitting complete model representations, knowledge transfer can reduce communication requirements while enabling useful information to be shared across distributed participants. Xue et al. (2022) similarly investigate two-timescale online gradient compression for over-the-air federated learning, illustrating how communication and optimization can be jointly designed.

Together, these studies establish that efficient edge-cloud AI requires multiple complementary mechanisms rather than a single compression technique.

2.3 Heterogeneity and Adaptive Intelligence

Edge devices rarely possess identical processing capabilities, memory capacities, network conditions, or model requirements. Hu et al. (2021) address this issue through Mhat, a model-heterogeneous aggregation training scheme for federated learning. This work supports the proposition that a distributed AI architecture must accommodate heterogeneous model structures rather than requiring every participant to execute an identical computational workload.

Zhang et al. (2022) further address heterogeneity through SplitAVG, a federated deep learning approach for medical imaging. Such architectures indicate that partitioning or restructuring model computation can be useful when participating environments differ substantially in available resources.

Reinforcement learning provides another mechanism for adaptive decision-making. Wen (2023) investigates federated offline reinforcement learning with multimodal data, indicating the potential for combining decentralized learning with complex decision environments. Xi (2024) focuses on reinforcement-learning-based real-time path planning for unmanned aerial vehicles, demonstrating the relevance of adaptive AI to latency-sensitive physical systems.

Yang et al. (2024), although focused on high-fidelity face-swapping, illustrate the computational demands of sophisticated deep learning architectures and therefore indirectly reinforce the need for scalable inference infrastructure.

The literature reveals a clear research gap: individual studies address communication efficiency, privacy, model heterogeneity, or adaptive intelligence, but an integrated edge-cloud framework must coordinate these dimensions simultaneously. Kumar (2025) provides a relevant architectural perspective by positioning inference as a resilient pipeline rather than an isolated model execution process. The present

framework extends this conceptual direction by connecting distributed inference with communication-aware learning and automated decision orchestration.

METHODOLOGY

3.1 Research Design

This study adopts a conceptual research-and-review methodology based exclusively on the supplied literature. Rather than conducting a new experimental benchmark, the methodology synthesizes the technical mechanisms reported in the selected studies and translates them into an integrated edge-cloud architecture. The analytical unit is the AI inference pipeline, comprising data acquisition, preprocessing, local inference, communication, model coordination, decision generation, and feedback.

The framework is constructed around five architectural principles: locality of computation, communication efficiency, heterogeneous participation, adaptive intelligence, and resilience. These principles are derived from the recurring technical problems identified across the literature.

3.2 Proposed Edge-Cloud AI Architecture

The proposed architecture consists of four functional layers: the edge sensing and inference layer, the communication optimization layer, the cloud intelligence layer, and the decision automation layer.

At the edge layer, devices collect data and perform preprocessing, feature extraction, and where feasible, direct AI inference. The principal objective is to minimize unnecessary data transmission. For example, a healthcare device can identify a preliminary anomaly locally and transmit only the relevant representation or model update rather than continuously transmitting raw physiological data.

The communication optimization layer determines what information should be transmitted and at what level of compression. Autoencoder-based compression, gradient compression, sparsification, adaptive rate control, and knowledge distillation represent complementary strategies. FLAC demonstrates how autoencoder compression can be integrated with federated learning (Beitollahi and Lu, 2022), whereas adaptive rate mechanisms address changing communication conditions (Cui et al., 2022). A resilient architecture should therefore dynamically select communication strategies according to bandwidth, latency, device capability, and model urgency.

The cloud intelligence layer provides large-scale model coordination, aggregation, historical analysis, and computationally intensive training. However, cloud coordination should not imply that every inference request must travel to the cloud. Instead, the cloud acts primarily as an intelligence coordination and model improvement layer, while edge nodes retain autonomy for latency-sensitive operations.

Model aggregation must account for heterogeneity. Mhat demonstrates the importance of model-heterogeneous aggregation (Hu et al., 2021), while SplitAVG demonstrates an alternative approach to distributed deep learning for heterogeneous medical imaging environments (Zhang et al., 2022). Therefore, the proposed framework allows different edge nodes to use appropriately scaled models while maintaining a common coordination mechanism.

The decision automation layer transforms model outputs into operational actions. This distinction is important because inference alone does not constitute automation. A real-time system must interpret predictions according to operational thresholds, confidence levels, resource conditions, and contextual constraints. Reinforcement learning can be used where decisions require sequential optimization, as demonstrated by its application to real-time path planning (Xi, 2024). In a hypothetical industrial setting, for example, an edge system could detect an abnormal machine condition, classify its severity locally, and automatically reduce operating speed while sending a summarized event to the cloud for long-term model refinement.

3.3 Adaptive Communication and Inference

A core methodological assumption is that edge-cloud communication should be adaptive. Let (L) denote inference latency, (C) communication cost, (E) computational energy, and (A) inference accuracy. A practical optimization objective can be conceptualized as:

[

\min ; F = \alpha L + \beta C + \gamma E - \delta A

]

where (\alpha,\beta,\gamma,\delta) represent application-specific priorities.

This formulation emphasizes that optimizing accuracy alone is insufficient. A model with marginally higher accuracy may be inappropriate if its execution or communication requirements cause unacceptable delays. Conversely, excessive compression may reduce communication cost but degrade model quality. The purpose of adaptive edge-cloud orchestration is therefore to identify a suitable operating point under changing system conditions.

SAFARI's focus on sparse communication under limited and unreliable connectivity supports this design principle (Mao, 2023). Similarly, over-the-air gradient compression demonstrates the potential for jointly exploiting communication characteristics and federated optimization (Xue et al., 2022). These findings support a framework in which communication policy becomes part of AI orchestration rather than an independent networking function.

3.4 Privacy, Security, and Trust

Privacy is another architectural requirement. Federated learning reduces the need for centralized raw-data collection, while healthcare-oriented research demonstrates how this principle can support sensitive distributed environments (Xu et al., 2021; Salim and Park, 2023). Blockchain integration offers an additional mechanism for distributed trust and accountability (Lakhan, 2023; Myrzashova et al., 2023).

Nevertheless, decentralized learning does not automatically eliminate security risks. Model updates may still expose information or become targets for manipulation. Consequently, the proposed framework treats privacy and security as cross-layer requirements rather than optional services. This approach is particularly important when automated decisions affect physical processes or sensitive individuals.

3.5 Resilience and Failure Handling

Resilience requires the architecture to remain operational despite unstable connectivity, heterogeneous nodes, and partial failures. Kumar (2025) emphasizes the importance of resilient edge-to-cloud inference pipelines for real-time decision systems. In the proposed architecture, resilience is achieved through local inference fallback, adaptive communication, asynchronous or partially synchronized model coordination, and cloud-edge workload redistribution.

For example, if a remote edge node loses cloud connectivity, it should continue executing its most recent validated model rather than becoming completely unavailable. Once connectivity returns, compressed updates can be synchronized. This design reduces dependence on continuous connectivity and aligns with the communication constraints identified in distributed federated learning research.

RESULTS

The synthesis produces five principal findings concerning robust edge-cloud AI computing.

First, communication efficiency is a primary determinant of real-time AI feasibility. Compression, sparsification, adaptive rate control, and knowledge distillation address different aspects of the communication problem (Beitollahi and Lu, 2022; Cui et al., 2022; Mao, 2023; Wu et al., 2022). The combined evidence suggests that no universal compression mechanism is optimal under all network

conditions. Consequently, adaptive communication policies are more appropriate than static transmission strategies.

Second, heterogeneity must be treated as an architectural property rather than an implementation anomaly. Edge nodes differ in computation, memory, connectivity, and model capability. Mhat and SplitAVG demonstrate that federated learning can be adapted to heterogeneous environments (Hu et al., 2021; Zhang et al., 2022). The proposed architecture therefore supports variable model complexity and flexible aggregation rather than enforcing uniform execution.

Third, federated learning provides an effective coordination mechanism for privacy-sensitive edge-cloud environments, particularly where raw data cannot easily be centralized. Healthcare applications demonstrate the practical relevance of this principle (Xu et al., 2021; Salim and Park, 2023). However, privacy must be integrated with security and trust mechanisms because decentralized computation alone does not guarantee complete protection.

Fourth, adaptive AI expands the role of edge-cloud infrastructure from prediction toward decision automation. Reinforcement learning applications demonstrate that distributed AI can support sequential and real-time decisions rather than merely producing static classifications (Wen, 2023; Xi, 2024). Kumar (2025) similarly positions the inference pipeline as a resilient mechanism for real-time decision systems. Thus, edge-cloud computing should be evaluated according to decision latency, reliability, and operational continuity in addition to predictive accuracy.

Fifth, resilience emerges from coordinated redundancy rather than from any single technology. Edge-local inference provides continuity during cloud disconnection, while cloud infrastructure provides computational scalability and global coordination. Communication compression reduces network dependency, and heterogeneous aggregation allows diverse devices to participate. The resulting architecture is therefore more robust than a purely centralized or purely decentralized design.

The findings also reveal an important trade-off. Increasing local computation can reduce communication latency but may exceed edge-device resource limits. Conversely, moving more computation to the cloud can improve computational capacity while increasing network dependency. The optimal architecture must therefore dynamically balance computation, communication, accuracy, and resilience.

DISCUSSION

The findings position edge-cloud AI computing as a multi-objective systems architecture rather than a simple extension of cloud AI. The theoretical contribution of the proposed framework lies in integrating five dimensions—computation locality, communication optimization, heterogeneity, adaptive intelligence, and resilience—into a single inference pipeline. Existing studies often optimize one or two of these dimensions, whereas real-time deployment requires their simultaneous consideration.

The first important implication concerns the relationship between latency and communication. Traditional cloud-centric inference assumes that data can be transmitted to centralized infrastructure without unacceptable delay. However, the communication-efficient federated learning literature indicates that network conditions can become a dominant constraint (Sun et al., 2020; Cui et al., 2022). The edge therefore becomes not merely a preprocessing location but an intelligent execution layer capable of producing decisions independently when necessary.

The second implication concerns model design. A single model architecture may be unsuitable for heterogeneous edge environments. Mhat and SplitAVG demonstrate the importance of accommodating heterogeneous model configurations (Hu et al., 2021; Zhang et al., 2022). This creates an architectural trade-off: model diversity improves deployment flexibility but complicates aggregation, validation, and lifecycle management. A robust system must therefore define compatibility mechanisms that allow different model configurations to contribute to a common intelligence layer.

The third implication concerns privacy and trust. Federated learning can reduce direct raw-data centralization, which is particularly valuable in healthcare (Xu et al., 2021). Blockchain-oriented studies further demonstrate the potential value of distributed trust mechanisms (Lakhan, 2023; Myrzashova et al., 2023). However, adding security mechanisms can increase computational and communication overhead. Thus, privacy protection itself becomes part of the resource-allocation problem.

A further theoretical implication concerns reinforcement learning. Wen (2023) and Xi (2024) illustrate that distributed AI can support complex sequential decision problems. In an edge-cloud architecture, reinforcement learning could be used not only for application-level decisions but also for infrastructure orchestration, such as determining whether a task should execute locally, be transmitted to the cloud, or be deferred. Such adaptive orchestration represents an important evolution beyond static edge-cloud workload placement.

The proposed framework also has practical limitations. First, the literature does not establish a universally optimal compression level because compression introduces potential accuracy and convergence trade-offs. Second, heterogeneous models complicate global coordination. Third, resilience mechanisms may preserve availability at the cost of temporarily operating with stale models. Fourth, security and blockchain mechanisms may impose additional overhead. Finally, real-world deployment requires empirical validation across different hardware, network conditions, and application domains.

Accordingly, the framework should not be interpreted as eliminating the edge-cloud trade-off. Instead, its contribution is to provide a structured mechanism for managing that trade-off dynamically. Kumar (2025) reinforces the significance of resilient inference pipelines, while the communication, federated learning, and reinforcement learning literature supplies complementary technical mechanisms. Future empirical research should evaluate the framework using latency, accuracy, communication volume, energy consumption, convergence behavior, failure recovery time, and decision reliability as joint performance metrics.

CONCLUSION

This study developed a robust conceptual framework for edge-cloud AI computing focused on real-time inference and decision automation. The analysis demonstrates that effective distributed intelligence requires more than relocating AI models between edge and cloud environments. It requires coordinated management of computation, communication, model heterogeneity, privacy, adaptive learning, and resilience.

The reviewed literature shows that federated learning can provide a foundation for decentralized model coordination, while compression, sparsification, adaptive rate control, and knowledge distillation can mitigate communication constraints. Heterogeneous aggregation mechanisms enable diverse edge devices to participate, while reinforcement learning supports dynamic decision-making. Healthcare-oriented studies further establish the importance of privacy, security, and trust in distributed AI environments.

The principal contribution of the proposed framework is the integration of these mechanisms into a unified inference pipeline. Edge devices provide low-latency local processing and operational continuity; communication mechanisms selectively transmit information; cloud infrastructure provides scalable coordination and model improvement; and decision automation converts inference outputs into operational actions. Such an architecture can support applications in healthcare, autonomous systems, industrial monitoring, smart infrastructure, and other latency-sensitive environments.

Nevertheless, several research challenges remain. Future work should empirically investigate dynamic workload placement, adaptive compression, heterogeneous model lifecycle management, secure aggregation, and failure-aware inference. Evaluation should move beyond isolated model accuracy toward multi-dimensional measures of latency, communication efficiency, energy consumption, reliability, privacy, and decision quality. The long-term research direction is therefore toward self-adaptive edge-cloud AI systems capable of continuously adjusting computational and communication strategies according to changing application and infrastructure conditions.

REFERENCES

1. M. Beitollahi and N. Lu, “FLAC: Federated learning with autoencoder compression and convergence guarantee,” in Proc. IEEE GLOBECOM Glob. Commun. Conf., 2022, pp. 4589–4594.

2. L. Cui, X. Su, Y. Zhou, and J. Liu, “Optimal rate adaption in federated learning with compressed communications,” in Proc. IEEE INFOCOM Conf, Comput. Commun., 2022, pp. 1459–1468.

3. L. Hu, H. Yan, L. Li, Z. Pan, X. Liu, and Z. Zhang, “Mhat: An efficient model-heterogenous aggregation training scheme for federated learning,” Inf. Sci., vol. 560, pp. 493–503, 2021.

4. A. Lakhan, “Federated-learning based privacy preservation and fraud-enabled blockchain IoMT system for healthcare,” IEEE J. Biomed. Health Inform., vol. 27, no. 2, pp. 664–672, Feb. 2023.

5. Y. Mao, “SAFARI: Sparsity-enabled federated learning with limited and unreliable communications,” IEEE Trans. Mobile Comput., early access, Jul. 18, 2023, doi: 10.1109/TMC.2023.3296624.

6. R. Myrzashova, S. H. Alsamhi, A. V. Shvetsov, A. Hawbani, and X. Wei, “Blockchain meets federated learning in healthcare: A systematic review with challenges and opportunities,” IEEE Internet Things J., vol. 10, no. 16, pp. 14418–14437, Aug. 2023.

7. M. M. Salim and J. H. Park, “Federated learning-based secure electronic health record sharing scheme in medical informatics,” IEEE J. Biomed. Health Inform., vol. 27, no. 2, pp. 617–624, Feb. 2023.

8. H. Sun, X. Ma, and R. Q. Hu, “Adaptive federated learning with gradient compression in uplink noma,” IEEE Trans. Veh. Technol., vol. 69, no. 12, pp. 16325–16329, Dec. 2020.

9. J. Wen, “Federated offline reinforcement learning with multimodal data,” IEEE Trans. Consum. Electron., early access, Nov. 08, 2023, doi: 10.1109/TCE.2023.3330943.

10. C. Wu, F. Wu, L. Lyu, Y. Huang, and X. Xie, “Communication-efficient federated learning via knowledge distillation,” Nature Commun., vol. 13, no. 1, 2022, Art. no. 2032.

11. M. Xi, “A lightweight reinforcement learning-based real-time path planning method for unmanned aerial vehicles,” IEEE Internet Things J., early access, Jan. 05, 2024, doi: 10.1109/JIOT.2024.3350525.

12. J. Xu, B. S. Glicksberg, C. Su, P. Walker, J. Bian, and F. Wang, “Federated learning for healthcare informatics,” J. Healthcare Inform. Res., vol. 5, pp. 1–19, 2021.

13. J. Yang, C. Cheng, S. Xiao, G. Lan, and J. Wen, “High fidelity face-swapping with style convtransformer and latent space selection,” IEEE Trans. Multimedia, vol. 26, pp. 3604–3615, 2024.

14. Y. Xue, L. Su, and V. K. Lau, “Fedocomp: Two-timescale online gradient compression for over-the-air federated learning,” IEEE Internet Things J., vol. 9, no. 19, pp. 19330–19345, Oct. 2022.

15. M. Zhang, L. Qu, P. Singh, J. Kalpathy-Cramer, and D. L. Rubin, “SplitAVG: A heterogeneity-aware federated deep learning method for medical imaging,” IEEE J. Biomed. Health Inform., vol. 26, no. 9, pp. 4635–4644, Sep. 2022.

16. K. Kumar, "Edge-to-Cloud AI Inference Pipelines: A Resilient Architecture for Real-Time Decision Systems," 2025 International Conference on Computer and Applications (ICCA), Bahrain, Bahrain, 2025, pp. 1-6, doi: 10.1109/ICCA66035.2025.11430905.

17. K. Ramamurthy, N. Bellamkonda and N. Amanmadov, "ScalePulse: Combinatorial Llm Framework for Scalability Constraint," SoutheastCon 2026, Huntsville, AL, USA, 2026, pp. 1-6, doi: 10.1109/SoutheastCon63549.2026.11476075

18. Philip, P. G. (2025). Explainable Artificial Intelligence (XAI) for Project Governance: Improving Transparency and Stakeholder Trust in Automated Project Decision. Journal of Project Management Studies, 2(1), 37–55. https://doi.org/10.58425/jpms.v2i1.570

19. Kodela, S., Kurada, S. B., Mogili, V. B., & Duggirala, J. (2026, March). Deep Learning-Enhanced Cloud Accounting Model for Real-Time Fraud and Financial Risk Prediction. In 2026 Innovations in Machine, Engineering, and Digital Conference (IMED) (pp. 1-6). IEEE