Journal of Emerging Intelligence and Engineering Technologies

Research Article

An Explainable Artificial Intelligence Framework for Improving Decision-Making Accuracy and Transparency in Complex Real-World Applications

Nurjonov Izzatbek[1], Suman Rana[2]

[1]Department of Accounting, Mamun University, Khiva, Uzbekistan

[2]Tulas University, Dehradun

Email: nurjanov.izzat@mamunedu.uz; Suman.rana@tulas.edu.in

Abstract

Complex decision-support systems increasingly rely on high-capacity machine learning, yet predictive performance alone cannot guarantee transparent, accountable or useful decisions. This paper proposes a modular explainable artificial intelligence (XAI) framework combining calibrated risk prediction, local and global feature attribution, counterfactual recourse, explanation-quality testing, drift surveillance and human review. A reproducible evaluation protocol is specified for heterogeneous real-world settings. To illustrate the planned analysis without claiming an executed experiment, numerical tables and charts report explicitly hypothetical demonstration values. The proposed architecture separates model accuracy from explanation fidelity and decision utility, allowing each to be tested independently. The paper identifies governance controls, operational limitations and falsifiable hypotheses for future empirical validation.

Keywords: Explainable artificial intelligence; interpretable machine learning; SHAP; LIME; counterfactual explanations; decision support; calibration; transparency; human oversight.

Received: 05 August 2025
Revised: 15 September 2025
Accepted: 10 October 2025
Published: 31 October 2025
DOI: 10.66592/fgd245se

© 2025 The Author(s). Published by the Journal of Emerging Intelligence and Engineering Technologies. This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0), which permits use, distribution, and reproduction in any medium for non-commercial purposes, provided the original work is properly cited. https://creativecommons.org/licenses/by-nc/4.0/

1. Introduction

Artificial intelligence has evolved from a collection of specialized prediction techniques into an infrastructure for supporting decisions in healthcare, finance, manufacturing, public administration and education. In these environments, a prediction may determine which case is reviewed first, which intervention is recommended, or which transaction receives additional scrutiny. A high classification score is valuable, but the real-world consequence depends on how predictions are translated into actions, what evidence is available to decision makers and whether errors are detected before harm occurs. Accordingly, decision-support quality must be understood as a combination of discrimination, calibration, explanation, workflow integration and governance.

2. Related Work

Ribeiro, Singh and Guestrin introduced LIME, a method that approximates the behavior of a classifier near a selected observation using a locally interpretable surrogate [1]. The approach is model agnostic and can be adapted to different input types. Its flexibility also creates methodological choices concerning neighborhood sampling, distance kernels and surrogate complexity. Because a

local explanation is fitted to perturbed examples, reproducibility and fidelity should be measured rather than presumed. The proposed framework therefore stores perturbation seeds and reports neighborhood agreement for each explanation.

3. Proposed Methodology

The proposed system consists of seven sequential stages: data governance and validation; preprocessing; candidate-model training; calibration and threshold selection; explanation and recourse generation; explanation-quality validation; and human-in-the-loop decision logging. Sensitive or immutable variables are screened before recourse generation. Every prediction is linked to model version, data schema, explanation parameters and audit metadata.

Data intake & consent →Validation / preprocessing →Train & calibrate models → Predict risk →SHAP / LIME / counterfactuals →Fidelity & fairness checks → Human review →Monitor drift

Let $D = \{(x_i, y_i)\}_{i=1}^{N}$, where $x_i \in \mathbb{R}^d$ and $y_i \in \{0,1\}$. The predictor $f\theta(x)$ estimates the probability of the positive class.

The predictive objective is $Lpred(\theta) = -(1/N)\ \Sigma_i\ [y_i \log f\theta(x_i) + (1-y_i) \log(1-f\theta(x_i))]$. This binary cross-entropy penalizes miscalibrated confident errors.

The decision rule is $\hat{a}_i = 1[f\theta(x_i) \geq \tau]$. The threshold τ is chosen using validation data and domain-specific false-positive and false-negative costs, not on the test set.

An additive explanation is written $f(x) \approx \varphi_0 + \Sigma_j\ \varphi_j(x)$, where φ_0 is the reference prediction and φ_j represents the attribution to feature j on a consistent output scale.

LIME solves $g^* = \text{argmin}(g \in G)\ \{L(f,g,\pi_x) + \Omega(g)\}$. Here π_x weights samples near x, L measures local surrogate disagreement and Ω controls explanation complexity.

A feasible counterfactual is $x^* = \text{argmin}(z \in F(x))\ [\lambda\ d(x,z) + \ell(f(z),y^*)]$, where F(x) enforces immutable-feature and domain constraints, d is a cost-weighted distance and y^* is the desired outcome.

Local fidelity is $\text{Fidelity}(x) = 1 - (1/M)\Sigma_k\ \ |f(z_k) - g(z_k)|$ for M neighborhood samples when model outputs lie in [0,1]. Larger values indicate closer surrogate agreement.

Decision accuracy is $Adec = (1/N)\Sigma_i\ 1[a_i = y_i]$, where a_i is the final human-system decision. This differs from stand-alone model accuracy and must be evaluated separately.

Calibration is evaluated with $Brier = (1/N)\Sigma_i\ (f\theta(x_i)-y_i)^2$; lower values are better. Subgroup false-negative rates and their uncertainty intervals are reported for governance review.

4. Experimental Design and Implementation Protocol

A future empirical study should use at least three datasets from distinct domains, for example publicly licensed health-risk, financial-risk and equipment-failure datasets, subject to appropriate access and ethics review. Data should be split into training (60%), validation (20%) and held-out testing (20%) by entity and time where relevant to prevent leakage. Missing values, encoding and scaling must be fitted on training data only. Compare logistic regression or a sparse rule-based model, random forest, gradient boosting, and the proposed calibrated explanation-enabled workflow. Use five random seeds or repeated temporal folds, confidence intervals and paired significance testing.

Suggested tools are Python, scikit-learn, SHAP, LIME, pandas and matplotlib. A human-factor experiment should randomly assign qualified reviewers to prediction-only, prediction-plus-explanation and transparent-model conditions. Measure reviewer correctness, time, inappropriate acceptance, explanation comprehension and confidence calibration. Pre-register endpoints and stratify by decision difficulty. Exclude protected variables from actionable recourse and conduct independent fairness review.

5. Results and Discussion — Illustrative Scenario Only

IMPORTANT: The following numerical values are hypothetical demonstration figures constructed to show the proposed reporting format. No model was trained and no participant study was conducted for this paper. The values must not be represented as observed empirical findings.

Table 1: Hypothetical comparative results for illustration; values are not measured.

Method	*Model accuracy %*	*Brier ↓*	*Local fidelity*	*Human decision accuracy %*
Interpretable baseline	84.2	0.114	0.91	79.4
Black-box predictor	88.6	0.092	0.0	81.0
Black-box + SHAP/LIME	88.6	0.092	0.87	85.1
Proposed calibrated XAI workflow	89.1	0.078	0.93	88.3

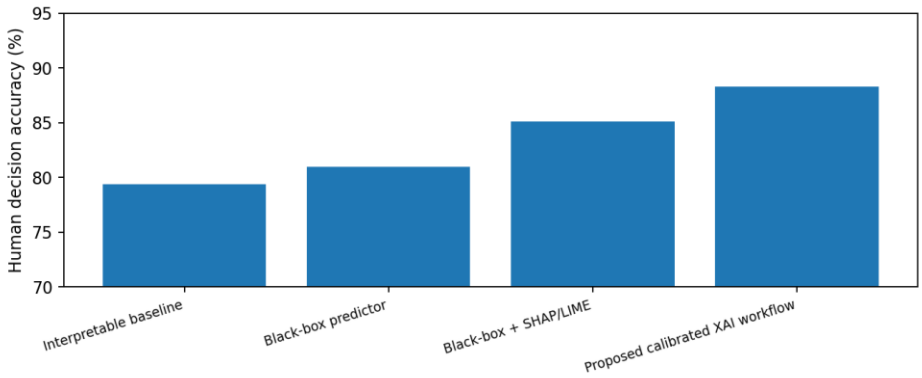


Figure 2: Hypothetical human decision accuracy comparison (not experimental evidence).

The illustrative table separates model accuracy from human decision accuracy. The unchanged model accuracy in the third row emphasizes that adding post-hoc explanations does not automatically improve a fixed predictor. The higher hypothetical human accuracy would require controlled user-study evidence, because additional context may improve some decisions while also increasing automation bias. Similarly, the proposed workflow's lower Brier score represents a possible effect of calibration and model selection, not a verified result. Fidelity values describe local surrogate agreement and are not interchangeable with fairness, causality or human comprehension.

A rigorous results section should report uncertainty intervals, dataset sizes, prevalence, subgroup confusion matrices, computational latency and the frequency of invalid counterfactuals. Sensitivity analysis should vary the

explanation neighborhood, SHAP background distribution, decision threshold and missingness rate. Failure cases must be documented alongside successful examples. An apparent gain is scientifically meaningful only if it generalizes to held-out conditions and survives paired statistical testing.

5.1. Threats to Validity and Robustness

Internal validity may be threatened by information leakage, uncontrolled reviewer expertise, and differences in the underlying case mix. All preprocessing must be nested within the training split, and reviewers should receive identical cases under randomized condition assignment. The comparison should distinguish errors attributable to the predictive model from those attributable to the reviewer's interpretation of its output. Confidence intervals should be computed at the participant and case level to reflect clustered observations.

Construct validity requires careful definitions of transparency and explanation utility. A high fidelity score does not guarantee that the explanation is understood; a user-friendly account does not establish that it reflects the model. A multi-dimensional evaluation should therefore include local surrogate error, perturbation stability, comprehension questions, time-to-decision, actionability of suggested changes, and a qualitative audit of misleading rationales. A preregistered composite index may be reported only alongside its individual components.

External validity is limited when all experiments use clean, retrospective benchmark datasets. Operational deployment often includes missing entries, shifting coding conventions, rare events and adversarial behavior. Future work should simulate progressive missingness, institution transfer and temporal distribution shift. Each scenario should be evaluated with the same fixed reporting protocol, including confidence intervals and documented failures. Evidence from one domain should not be generalized to another without independent validation.

Explanation robustness can be assessed by perturbing irrelevant features and measuring rank agreement among the most important features. However, perturbations must remain within plausible data ranges; otherwise instability may simply reflect unrealistic inputs. Counterfactual quality likewise requires feasibility checks, sparsity, plausibility and diversity, rather than distance alone. Domain experts should determine which variables are genuinely actionable and whether suggested changes could produce unintended consequences.

Governance considerations include auditability, access control, data minimization, and the risk of revealing sensitive attributes through explanation outputs. Explanations should be tailored to the role of the viewer. Technical staff may need full feature-level diagnostics, while frontline users may need concise reasons, confidence warnings and escalation guidance. Access to personally identifiable features and sensitive proxies should be restricted according to applicable policies and law.

5.2. Deployment and Practical Implications

A production implementation should treat every prediction as a versioned event. The log should include timestamp, feature schema version, model identifier, calibration version, explanation method and background sample identifier. This metadata supports reproducibility and investigation of inconsistent decisions. Monitoring should trigger review when calibration,

subgroup performance, missingness or explanation stability crosses predefined limits. Such thresholds should be validated against operational risk tolerance.

Human oversight should be designed as a specific decision protocol rather than a generic instruction to review outputs. Cases can be routed to manual review when uncertainty is high, an explanation fails fidelity checks, or a counterfactual violates domain constraints. Reviewers should be able to record reasons for overrides and flag explanations that conflict with known domain relationships. These records can support future quality improvement without treating every override as ground truth.

The framework also implies a resource trade-off. Computing local explanations and multiple counterfactual candidates can increase latency and infrastructure cost. A practical implementation may cache global summaries, compute detailed local explanations only for borderline decisions, and use interpretable models when comparable predictive performance is achieved. Cost, response time and energy use should be reported beside predictive and human-outcome metrics.

Finally, an effective deployment strategy should establish a clear boundary between prediction and recommendation. The model estimates risk under observed historical conditions; it does not automatically establish the causal effect of a proposed intervention. Users must be told when a counterfactual is a mathematical recourse suggestion rather than a validated intervention. Prospective evaluation, expert approval and post-deployment incident review are necessary before asserting that the system improves real-world outcomes.

6. Conclusion

This paper presents a modular XAI decision-support framework designed to integrate predictive reliability, explanation fidelity, feasible counterfactuals and accountable human oversight. The main methodological finding is that predictive accuracy, transparency and decision quality require separate measurement; none can safely be inferred from the others. The proposed architecture offers a reproducible protocol for comparing interpretable baselines, opaque models and explanation-enabled workflows while monitoring drift, calibration and subgroup performance. The numerical illustration demonstrates how results may be reported but does not establish empirical superiority. Future work should implement the protocol across multiple domains, conduct randomized human evaluations and assess explanation stability, latency and equitable outcomes before any deployment claims are made.

References

Kumar, Ashok, and Arpit Jain. "Image smog restoration using oblique gradient profile prior and energy minimization." Frontiers of Computer Science 15, no. 6 (2021): 156706.

Rangappa, A. S. (2022). The use of blockchain for auditable and tamper-proof database security logs: Enhancing transparency and trust in data access monitoring. International Journal on Recent and Innovation Trends in Computing and Communication, 10(9), 254–261. https://www.ijritcc.org/index.php/ijritcc/article/view/11925

Anchala, S. (2023). Utilization of blockchain technology in cloud storage systems for enhancing data integrity and preventing unauthorized modifications through decentralized validation. International Journal on Recent and Innovation Trends in Computing and Communication, 11(5), 629–637. https://www.ijritcc.org/index.php/ijritcc/article/view/11942

Chatrath, A. (2024). The role of encryption and key management in cloud security: Balancing performance, scalability, and confidentiality in multi-tenant and distributed data storage systems. International Journal of Innovative Research in Science, Engineering and Technology, 13(7). https://doi.org/10.15680/IJIRSET.2024.1307187

Devulapalli, N. (2019). Post-quantum cryptography for database security: Preparing for the impact of quantum computing on encrypted data storage and secure queries. Journal of Critical Reviews, 6(1), 614–620. https://doi.org/10.69980/jcr.v6:i1.13396

Kumar, S., Jain, A., & Moparthi, N. R. (2022). Enhanced method of object tracing using extended Kalman filter via binary search algorithm.

Talasila, D. (2020). Deployment of shift-left security practices in agile development for embedding security controls early in the lifecycle and reducing technical debt through dev-centric training. Turkish Online Journal of Qualitative Inquiry, 11(1), 643–651. https://doi.org/10.69980/jzqq0t93

Devulapalli, N. (2023). Database security in disaster recovery and business continuity planning: Ensuring data availability, backup integrity, and secure failover mechanisms. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3), 820–827. https://www.ijritcc.org/index.php/ijritcc/article/view/11944

Chatrath, A. (2021). Implementation of AI in recruitment platforms for improving candidate screening and diversity inclusion through bias-free resume parsing and skill matching. International Journal of Innovative Research in Computer and Communication Engineering, 9(12). https://doi.org/10.15680/IJIRCCE.2021.0912046

Ahuja, R. (2023). Implementation of AI-driven input validation in user interfaces for preventing injection attacks and ensuring data sanitation through adaptive filtering and pattern recognition. International Journal on Recent and Innovation Trends in Computing and Communication, 11(8), 884–891. https://www.ijritcc.org/index.php/ijritcc/article/view/11936

Mahadasu, S. S. (2020). Cloud-native life cycle management of virtualized RAN functions using Kubernetes operators and policy-driven automation. International Journal of Engineering Science and Advanced Technology, 20(4), 21–26. https://doi.org/10.64771/ijesat.2020.v20.i04.pp3846

S. Goyal, K. K. Gattupalli, B. Rallabandi and M. S. Medarametla, "Integrating Terrestrial and Non-Terrestrial Networks for Reliable Rural Coverage in Agriculture and Smart Grids," 2025 UPWIECON, Dehradun, India, 2025, pp. 846–850, doi: 10.1109/UPWIECON67212.2025.11390331.

Bhanushali, B. (2024). Artificial intelligence-powered anti-money laundering framework for enhancing suspicious transaction detection, automating compliance processes, and strengthening fraud prevention in financial institutions. International Journal of Multidisciplinary Research in Science, Engineering and Technology, 7(10), 16126–16134. https://doi.org/10.15680/IJMRSET.2024.0710080

Jain, Arpit, Adesh Kumar, and Sanjeev Sharma. "Comparative design and analysis of mesh, torus and ring NoC." Procedia Computer Science 48 (2015): 330–337.

Aggarwal, A. (2021). Implementation of machine learning in cloud SIEM solutions for enhancing threat detection accuracy and reducing alert fatigue through correlation-based anomaly detection. NeuroQuantology, 19(10). https://doi.org/10.48047/nq.2021.19.10.NQ21226

N. Tripathi, K. K. Gattupalli, B. Rallabandi and M. S. Medarametla, "AI-Native Cloud-RAN Orchestration for Enterprise Private 5G Using Digital Twin Models," 2025 1st IEEE UP Section WIE Int. Conf. on Electrical Electronics and Computer Engineering (UPWIECON), Dehradun, India, 2025, pp. 851–857, doi: 10.1109/UPWIECON67212.2025.11390348.

Gupta, A. (2024). Containerization, virtualization, and cloud security: Analyzing vulnerabilities in Kubernetes, Docker, and hypervisor-based architectures with mitigation strategies. International Journal of Advanced Research in Education and Technology, 11(3), 1322–1331. https://doi.org/10.15680/IJARETY.2024.1103067

K. M. Zubair, B. R. Rallabandi, A. Yadav, K. Kaushik, S. Wankhade and M. Soni, "Explainable AI-Driven Intrusion Response Architecture for Autonomous Cyber

Defense Systems," 2025 Int. Conf. on Digital Innovations for Sustainable Solutions (ICDISS), Faridabad, India, 2025, pp. 1–6, doi: 10.1109/ICDISS68238.2025.11320639.

Ahuja, R. (2022). Application of static application security testing (SAST) in CI/CD pipelines for early detection of insecure coding practices through syntax tree analysis and custom rule sets. International Journal on Recent and Innovation Trends in Computing and Communication, 10(2), 80–87. https://www.ijritcc.org/index.php/ijritcc/article/view/11922

Gupta, A. (2021). Deployment of advanced endpoint detection in remote work environments for preventing malware intrusions and data exfiltration through behavioral analytics and AI-driven responses. NeuroQuantology, 19(9), 1206–1224. https://doi.org/10.48047/nq.2021.19.9.NQ21199

Jain, Arpit, Nageswara Rao Moparthi, A. Swathi, Yogesh Kumar Sharma, Nitin Mittal, Ahmed Alhussen, Zamil S. Alzamil, and MohdAnul Haq. "Deep Learning-Based Mask Identification System Using ResNet Transfer Learning Architecture." Computer Systems Science & Engineering 48, no. 2 (2024): 341.

Mahajan, R. A., Dey, R., Khan, M., Kanabar, K., & Gupta, S. K. (2025). Real-time vehicular-to-vehicular (V2V) communication and applications for pedestrian safety in AI-optimized VANETs for autonomous vehicles. In H. Maryam, M. M. Malik, I. U. Khan, & S. K. Gupta (Eds.), AI-driven transportation systems: Real-time applications and related technologies (pp. 151–167). Springer. https://doi.org/10.1007/978-3-031-98349-8_8

Talasila, D. (2024). Adoption of data loss prevention strategies in SaaS platforms for safeguarding confidential information and preventing unauthorized transfers through real-time content inspection. International Journal of Multidisciplinary Research in Science, Engineering and Technology, 7(7), 13537–13545. https://doi.org/10.15680/IJMRSET.2024.0707293

Khan, S. (2017). The role of privacy-preserving techniques in database security: Secure multi-party computation, differential privacy, and homomorphic encryption. The Research Journal, 3(4), 29–35.

Dwivedi, D. (2025). A comparative study of public, private, and hybrid cloud security architectures: Evaluating trade-offs in scalability, privacy, cost, and risk management. International Journal of Multidisciplinary Research in Science, Engineering and Technology, 8(1), 485–493. https://doi.org/10.15680/IJMRSET.2025.0801069

Aggarwal, A. (2022). Cloud security in artificial intelligence workloads: Protecting model training, data confidentiality, and AI-generated insights in cloud-based machine learning pipelines. NeuroQuantology, 20(3), 1367–1382. https://doi.org/10.48047/nq.2022.20.3.NQ23004

Gupta, A. (2022). The role of artificial intelligence in transforming healthcare delivery: A critical analysis of machine learning models, predictive diagnostics, and ethical challenges in medical decision-making. NeuroQuantology, 20(3), 1383–1399. https://doi.org/10.48047/nq.2022.20.3.NQ23005

Mahadasu, S. S. (2021). Telemetry-assisted mobility management for 5G standalone networks with distributed edge user-plane functions. International Journal of Engineering Research and Science & Technology, 17(1), 141–145. https://doi.org/10.62643/ijerst.2019.v17.n1.4406

Jain, A. (2024, December). Energy-Efficient Signal Processing Architectures for IoT Networks Design. In 2024 International Conference on Emerging Technologies and Innovation for Sustainability (EmergIN) (pp. 362–367). IEEE.

Aggarwal, A. (2024). Integration of serverless security mechanisms in Function-as-a-Service platforms for preventing execution abuse and data leakage through contextual isolation techniques. International Journal of Innovative Research in Computer and Communication Engineering, 12(5). https://doi.org/10.15680/IJIRCCE.2024.1205374

Talasila, D. (2023). Implementation of cloud-based disaster recovery models for minimizing business downtime and ensuring operational continuity through geographically distributed backup systems. International Journal on Recent and Innovation Trends in Computing and Communication, 11(8), 867–874. https://www.ijritcc.org/index.php/ijritcc/article/view/11934

Bhanushali, B. (2025). Leveraging block chain and zero-knowledge proofs for enhancing privacy in financial transactions, securing sensitive data, and mitigating cyber security risks in digital payment systems. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 14(1), 130–138. https://doi.org/10.15662/IJAREEIE.2025.1401020

Anchala, S. (2023). Integration of confidential computing in multi-cloud environments for enhancing data privacy and regulatory compliance through encrypted processing and isolated execution. International Journal on Recent and Innovation Trends in Computing and Communication, 11(9), 5698–5709. https://www.ijritcc.org/index.php/ijritcc/article/view/11933

Chatrath, A. (2025). Integration of AI in air quality monitoring systems for enhancing environmental health and public awareness through predictive analytics and real-time sensing networks. International Journal of Multidisciplinary Research in Science, Engineering and Technology, 8(1). https://doi.org/10.15680/IJMRSET.2025.0801068

Ahuja, R. (2025). The risks and opportunities of generative AI in cyber security: Investigating the dual-use nature of generative models for phishing, malware development, and defence through synthetic data and automated threat detection. International Journal of Innovative Research in Computer and Communication Engineering, 13(2), 1758–1767. https://doi.org/10.15680/IJIRCCE.2025.1302059

Jain, Jaishree, Arpit Jain, Saurabh Kumar Srivastava, Chaman Verma, Maria Simona Raboaca, and Zoltán Illés. "Improved security of e-healthcare images using hybridized robust zero-watermarking and hyper-chaotic system along with RSA." Mathematics 10, no. 7 (2022): 1071.

Rangappa, A. S. (2024). Generative AI in human-machine collaboration: Exploring co-creation dynamics, shared agency, and the role of human oversight in AI-augmented creative and analytical tasks. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 13(4), 985–992. https://doi.org/10.15662/IJAREEIE.2024.1304077

Anchala, S. (2024). Deployment of AI-enabled surveillance systems in urban areas for strengthening public safety and crime prevention through real-time facial recognition and anomaly detection. International Journal of Multidisciplinary Research in Science, Engineering and Technology, 7(10), 16116–16125. https://doi.org/10.15680/IJMRSET.2024.0710079

Aggarwal, A. (2020). Deployment of multi-factor authentication in cloud-based enterprise applications for preventing credential theft and unauthorized access through adaptive verification layers. International Journal of Multidisciplinary Research in Science, Engineering and Technology, 3(8). https://doi.org/10.15680/IJMRSET.2020.0308009

Dwivedi, D. (2024). Agentic AI and global governance: Addressing the challenges of regulating autonomous artificial agents across borders in a digitally interconnected world. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 13(11), 2527–2534. https://doi.org/10.15662/IJAREEIE.2024.1311011

Bhanushali, B. (2020). Enhancing alert investigation process with feedback loop to increase efficiently and model tuning. International Journal of Innovative Research in Computer and Communication Engineering, 8(8), 3366–3373. https://doi.org/10.15680/IJIRCCE.2020.0808100

Mahadasu, S. S. (2019). Performance-aware VNF placement for distributed 5G core networks over heterogeneous edge and cloud infrastructure. International Journal of Engineering Science and Advanced Technology, 19(10), 5–11. https://doi.org/10.64771/ijesat.2019.v19.i10.pp3849

Rao, S. Madhusudhana, and Arpit Jain. "Advances in Malware Analysis and Detection in Cloud Computing Environments: A Review." International Journal of Safety & Security Engineering 14, no. 1 (2024).

Ahuja, R. (2024). Deployment of content security policies in front-end frameworks for blocking cross-site scripting and clickjacking through declarative access restrictions. International Journal of Multidisciplinary Research in Science, Engineering and Technology, 7(11), 17234–17242. https://doi.org/10.15680/IJMRSET.2024.0711132

S. Goyal, K. K. Gattupalli, B. Rallabandi and M. S. Medarametla, "Intent-Driven Zero-Touch Provisioning in Private 5G Networks with AI/ML Automation," 2025 UPWIECON, Dehradun, India, 2025, pp. 839–845, doi: 10.1109/UPWIECON67212.2025.11390129.

Chatrath, A. (2021). Bias, fairness, and ethics in artificial intelligence: A comprehensive study of algorithmic discrimination, transparency challenges, and governance frameworks. International Journal of Multidisciplinary Research in Science, Engineering and Technology, 4(1). https://doi.org/10.15680/IJMRSET.2021.0401014

Khan, S. (2016). A study of data provenance and integrity in database security: Ensuring authenticity, non-repudiation, and accountability in data lifecycle management. International Journal of Research in Electronics and Computer Engineering, 4(3), 180–186.

Mahajan, R. A., Dey, R., Khan, M., Mohd Su'ud, M., Alam, M. M., & Jadhav, P. (2025). Enhancing personalization in IoT-based health monitoring via generative AI and transfer learning. Egyptian Informatics Journal, 32, Article 100788. https://doi.org/10.1016/j.eij.2025.100788

Harshitha, G., Kumar, S., Rani, S., & Jain, A. (2021, November). Cotton disease detection based on deep learning techniques. In 4th Smart Cities Symposium (SCS 2021) (Vol. 2021, pp. 496–501). IET.

Bhanushali, B. (2024). Tokenization of real-world assets through block chain for improving liquidity, enabling fractional ownership, and reducing transaction costs in global financial markets. International Journal of Advanced Research in Education and Technology, 11(4), 1809–1816. https://doi.org/10.15680/IJARETY.2024.1104071

K. Cherladine, B. R. Rallabandi, A. Goel, K. Kaushik, A. Dhillon and M. Soni, "Generative Adversarial Defense Models for Simulated Cyberattack Scenarios in Virtual Networks," 2025 ICDISS, Faridabad, India, 2025, pp. 1–6, doi: 10.1109/ICDISS68238.2025.11320686.

Mahadasu, S. S. (2022). GitOps-based continuous delivery of cloud-native 5G network functions with automated configuration validation. International Journal of Data Science and IoT Management System, 1(1), 50–56. https://doi.org/10.64751/ijdim.2022.v1.n1.1286

Kumar, S., Sharma, K., Kumar, P. A., Jain, A., Bhagat, S. K., & Singh, P. (2024, September). An improved particle swarm approach for energy-aware location-aided routing in mobile ad-hoc network. In 2024 7th International Conference on Contemporary Computing and Informatics (IC3I) (Vol. 7, pp. 1119–1124). IEEE.

Dwivedi, D. (2024). Deployment of AI-based CAPTCHA systems in web applications for preventing automated bot attacks and brute force attempts through adaptive human verification. International Journal of Multidisciplinary Research in Science, Engineering and Technology, 7(7), 13529–13536. https://doi.org/10.15680/IJMRSET.2024.0707292

Talasila, D. (2021). A risk-based framework for database security: Identifying, classifying, and prioritizing threats for adaptive security policy enforcement. International Journal on Recent and Innovation Trends in Computing and Communication, 9(1), 54–61. https://www.ijritcc.org/index.php/ijritcc/article/view/11919

Devulapalli, N. (2020). Implementation of compliance-as-code in regulated DevSecOps environments for automating audit readiness and reducing human error through scripted control validation. Journal of Critical Reviews, 7(1), 2651–2657. https://doi.org/10.69980/jcr.v7:i1.13405

Dwivedi, D. (2018). DevSecOps and API security: Embedding secure development practices, automated testing, and continuous monitoring in API-centric architectures. International Journal of Innovative Research in Computer and Communication Engineering, 6(1), 511–519. https://doi.org/10.15680/IJIRCCE.2018.0601084

Khan, S. (2022). Sales force security in mobile and remote work environments: Addressing authentication challenges, secure synchronization, and device-level threats. International Journal of Innovative Research in Computer and Communication Engineering, 10(12), 8735–8743. https://doi.org/10.15680/IJIRCCE.2022.1012040

Rangappa, A. S. (2023). An analytical study of role-based access control (RBAC) and attribute-based access control (ABAC) in database security for multi-tenant and cloud-based architectures. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3), 812–819. https://www.ijritcc.org/index.php/ijritcc/article/view/11943

Gupta, Namit, Kunwar Singh Vaisla, Arpit Jain, Adesh Kumar, and Rajeev Kumar. "Performance Analysis of AODV Routing for Wireless Sensor Network in FPGA Hardware." Computer Systems Science & Engineering 40, no. 3 (2022).