International Journal of Research and Applied Innovations (IJRAI)

| ISSN: 2455-1864 | www.ijrai.org | editor@ijrai.org | A Bimonthly, Scholarly and Peer-Reviewed Journal |

||Volume 8, Issue 6, November–December 2025||

DOI:10.15662/IJRAI.2025.0806016

Explainable Generative ML–Driven Cloud-Native Risk Modeling with SAP HANA–Apache Integration for Data Safety

R.Sugumar

Institute of CSE, SIMATS Engineering, Chennai, India

ABSTRACT: Financial institutions increasingly require predictive models that are not only accurate but auditable, fair, and deployable at enterprise scale. This paper proposes an architectural and methodological framework for Cloud-Native Explainable Generative Machine Learning (XGen-ML) tailored to credit scoring and risk modeling, designed to run across SAP HANA (in-memory analytical platform) and the Apache ecosystem (Spark, Kafka, Hadoop) for large-scale ingestion, feature engineering, model training, synthetic data generation, and real-time scoring. The approach integrates three technical pillars: (1) generative modeling (GANs/VAEs/conditional generative models) to produce privacy-preserving synthetic tabular datasets for model training, stress-testing, and imbalance mitigation; (2) explainable AI (XAI) techniques — combining global explanation methods and instance-level interpretability (e.g., SHAP, LIME, rule extraction) — to provide audit trails for regulatory compliance and to expose drivers of decisions to business users; and (3) a cloud-native serving and governance layer using containerized microservices, model registries, feature stores, and stream processing to enable continuous training, monitoring, and model risk management. Generative models help address the scarcity and privacy constraints of sensitive credit datasets and can augment minority classes to reduce biased model performance, while XAI techniques ensure transparency, support human review, and enable contestability of decisions. The proposed system leverages SAP HANA's in-memory SQL engine for high-performance analytical joins and feature computation, and the Apache stack (Spark for distributed ETL/ML, Kafka for streaming, Hadoop for archival) for scalable data processing and orchestration. We describe a reproducible methodological pipeline: (i) secure data ingestion and schema harmonization; (ii) synthetic data generation and augmentation using conditional tabular GANs; (iii) feature engineering and representation learning in Spark with push-down analytics to HANA where appropriate; (iv) model training using hybrid ensembles (tree-based + neural generative components) and causal feature selectors; (v) local and global explainability modules (SHAP for global attributions, LIME for local explanations, surrogate rules for regulatory reporting); (vi) deployment to a cloud-native inference service with runtime logging for model governance; and (vii) continuous monitoring for drift, fairness metrics, and performance degradation. We evaluate the framework qualitatively against regulatory requirements (e.g., explainability, auditability), operational constraints (latency, throughput), and model quality metrics (AUC, calibration, false-positive/negative cost asymmetries). Drawing on literature and applied case studies, we show that combining generative approaches with rigorous XAI and production-grade data platforms reduces the trade-off between performance and transparency while improving robustness to dataset shifts. The paper concludes with practical considerations for privacy (differential privacy extensions for generative models), governance (model registries, explainability documentation), limitations (synthetic data fidelity, worst-case fairness scenarios), and future extensions (causal generative models, privacy-preserving federated training). Key contributions are an end-to-end cloud-native architecture and a prescriptive methodology for integrating generative ML and XAI into credit and risk pipelines suitable for enterprise SAP HANA + Apache environments.

KEYWORDS: Cloud-native; Explainable AI (XAI); generative models; GAN; CTGAN; SHAP; LIME; credit scoring; credit risk; SAP HANA; Apache Spark; synthetic data; model governance; fairness; feature store.

I. INTRODUCTION

Credit and risk modeling is a cornerstone of modern financial services: lenders, card issuers, and regulators all depend on predictive models to estimate default probabilities, set limits, price risk, and satisfy compliance obligations. Traditional scorecard models (e.g., logistic regression and manual scorecards) emphasize interpretability, regulatory tractability, and simple business rules, but can lack the predictive power required in complex, high-dimensional datasets. In contrast, modern machine learning (ML) methods — tree ensembles and deep neural networks —

| ISSN: 2455-1864 | www.ijrai.org | editor@ijrai.org | A Bimonthly, Scholarly and Peer-Reviewed Journal |

||Volume 8, Issue 6, November–December 2025||

DOI:10.15662/IJRAI.2025.0806016

frequently achieve better out-of-sample performance but are often criticized as “black boxes,” making regulatory explainability, dispute resolution, and fairness assessment more difficult. This tension has driven two complementary trends that motivate this work: (1) the rise of explainable AI (XAI) methods that produce local and global explanations for complex models, and (2) the application of generative ML to create synthetic, privacy-preserving, or class-balanced tabular data useful for training, robustness testing, and augmentation.

Explainability matters in credit: regulators and internal risk committees require audit trails, feature-level explanations, and evidence that adverse action notices are supported by understandable reasons. Robust explainability techniques now exist that map complex model behavior into additive feature attributions (SHAP) or create local surrogate models to explain single decisions (LIME), enabling a plausible route toward reconciling model complexity with accountability. The idea of combining XAI with generative methods is powerful: synthetic data can enable richer model training without exposing customer PII, and XAI can provide interpretability for decisions made by models trained on real and synthetic mixtures. SHAP and LIME are two widely used tools that have been validated across many financial use-cases and are central to our proposed explainability stack. (See Lundberg & Lee, 2017; Ribeiro et al., 2016.) (arXiv)

Generative modeling for tabular data has matured rapidly: conditional and specialized GAN variants for mixed continuous/discrete columns (e.g., CTGAN) have demonstrated the ability to reproduce complex joint distributions for tabular datasets and to assist with rare-class synthesis, which is particularly relevant for credit defaults that are typically imbalanced. Empirical studies have shown CTGAN and later CTAB-GAN variants to outperform simpler oversampling approaches under many conditions, though careful fidelity assessment and downstream task validation remain necessary. (See Xu et al., 2019; Zhao et al., 2021.) (arXiv)

From an engineering and operations perspective, enterprises rarely deploy models as one-off experiments; they require robust data platforms capable of high throughput, low latency scoring, and rigorous governance. SAP HANA’s in-memory SQL engine is already used in many financial institutions for high-performance analytics and operational decisioning, and it can integrate with Apache Hadoop/Spark for scale and with Kafka for streaming. Leveraging this hybrid architecture lets teams push compute to HANA for fast joins, while using Spark for distributed ETL/feature pipelines and model training on clusters. Documented integration patterns between SAP HANA and the Apache stack make this hybrid design practical for production. (SAP Help Portal)

This paper describes an end-to-end Cloud-Native XGen-ML framework for credit risk: it couples generative synthetic data pipelines, explainability modules, and cloud-native microservices running over SAP HANA and the Apache ecosystem. We present a prescriptive research methodology (detailed later) and discuss operational considerations — latency targets, audit logging, privacy constraints, model registries, and fairness monitoring. We also present a discussion of the inherent limitations of synthetic data fidelity and explainability approximations and propose practical mitigations (privacy-aware GANs, extensive back-testing, and human-in-the-loop review). The intended audience is data science teams in banks/fintechs, ML engineers responsible for model ops, and regulators seeking practical architectures that balance predictive power with transparency.

II. LITERATURE REVIEW

Research on credit scoring and machine learning has a long history. Early large-scale evaluations showed that ML methods could significantly outperform linear models in credit prediction tasks when plentiful transactional and bureau data were available (Khandani, Kim & Lo, 2010). That work helped catalyze the industry turn toward ensemble methods and paved the way for adoption of tree ensembles and gradient boosted machines in credit risk applications. (ScienceDirect)

Explainability methods have emerged to mitigate the trade-off between model complexity and interpretability. LIME (Ribeiro et al., 2016) introduced local surrogate explanations, enabling practitioners to inspect the model’s local decision boundary for a single instance; SHAP (Lundberg & Lee, 2017) provided a unified, game-theoretic attribution framework that has become a de-facto standard for global and local feature contributions. These methods are widely recommended in financial model governance because they provide consistent, additive attributions and can be adapted to produce counterfactual reasoning or compact rule-like explanations appropriate for regulatory reporting. (arXiv)

International Journal of Research and Applied Innovations (IJRAI)

| ISSN: 2455-1864 | www.ijrai.org | editor@ijrai.org | A Bimonthly, Scholarly and Peer-Reviewed Journal |

||Volume 8, Issue 6, November–December 2025||

DOI:10.15662/IJRAI.2025.0806016

Generative modeling for tabular data (GANs, VAEs, and hybrids) has matured rapidly from early proofs-of-concept to practical tools. The CTGAN family (Xu et al., 2019) and subsequent CTAB-GAN and CTAB-GAN++ variants were designed to address mixed discrete/continuous columns and imbalanced modes, improving downstream classifier performance when synthetic data is used for training or augmentation. Studies show generative models often outperform classical oversampling (SMOTE) in reproducing complex joint distributions, but they require careful evaluation against statistical fidelity metrics and downstream task performance measures. (arXiv)

Recent work has specifically evaluated generative models for financial use cases. Papers in 2023–2024 examined synthetic financial series for VaR estimation, GAN-based oversampling for rare defaults, and hybrid autoencoder-GAN pipelines for credit data augmentation; these studies emphasize both the practical benefits (data augmentation, privacy) and the pitfalls (mode collapse, loss of subtle dependencies). Robust assessment frameworks that compare synthetic and real datasets using inter-column dependency metrics and downstream task agreement have been proposed and are recommended before synthetic data enters production pipelines. (Taylor & Francis Online)

Platform integration matters. SAP HANA documentation and technical guides explain patterns for integrating HANA with Hadoop/Spark for hybrid workloads: HANA excels at low-latency, in-memory analytical joins and can act as a feature store or operational DB, while Spark/Hadoop provide distributed ETL, model training, and archival storage. Industry case studies show successful real-time credit use cases built on Spark MLlib with push-down analytics to HANA for feature scoring (architectural patterns include virtual tables, smart data access, and Spark-to-HANA connectors). (SAP Help Portal)

The convergence of XAI and generative ML raises several research and operational questions: How to ensure synthetic data does not introduce spurious dependencies that mislead models? How to calibrate XAI outputs for ensemble models used in scoring? How to provide legally defensible explanations for automated decisions? The literature provides partial answers — use downstream validation, incorporate fairness constraints, and adopt model governance practices — but practical end-to-end blueprints remain scarce. Our framework addresses these gaps by combining synthesis fidelity checks, XAI-backed model documentation, and cloud-native deployment patterns to make XGen-ML feasible in financial production settings. Representative surveys and comparative studies on generative model assessment and XAI evaluation guide our methodology. (ScienceDirect)

III. RESEARCH METHODOLOGY

Below I present the proposed research methodology as a sequence of repeatable, list-style paragraphs (each paragraph names a stage, describes inputs/outputs, methods, evaluation criteria, and practical notes).

1. **Data Inventory, Access Controls, and Privacy Scoping.**

o Inputs: raw source schemas (transactional systems, credit bureaus, application forms), PII catalog, regulatory requirements.

o Process: create a data inventory and data lineage map; classify fields by sensitivity; define anonymization and retention policies. Establish secure connectors to source systems (Kafka for event streams, JDBC for HANA/OLTP sources). Map required features for model prototypes and for production scoring.

o Outputs: an approved data access matrix, data dictionary, and a privacy impact assessment (PIA).

o Evaluation: review by legal/compliance, ensure ability to produce adverse action explanations from available fields.

o Notes: early PIA reduces rework during model governance.

2. **Schema Harmonization and Feature Computation (Spark + Push-down to HANA).**

o Inputs: lineage map from Step 1; raw streams and batch dumps.

o Process: canonicalize entity IDs, time windows, and variable encodings. Execute distributed ETL in Spark to compute expensive aggregations, with smaller, latency-sensitive joins pushed down to SAP HANA via virtual tables where applicable to exploit HANA's in-memory performance. Implement a feature store (cataloging feature definitions, freshness metadata, and transformation code).

o Outputs: versioned feature artefacts in the feature store; training and validation datasets.

o Evaluation: consistency checks (row counts, null rates), schema drift detectors.

o Notes: pushing suitable operations to HANA reduces network I/O and lowers online scoring latency.

International Journal of Research and Applied Innovations (IJRAI)

| ISSN: 2455-1864 | www.ijrai.org | editor@ijrai.org | A Bimonthly, Scholarly and Peer-Reviewed Journal |

||Volume 8, Issue 6, November–December 2025||

DOI:10.15662/IJRAI.2025.0806016

3. **Exploratory Data Analysis and Imbalance Diagnosis.**

o Inputs: training datasets.

o Process: quantify class imbalance (delinquency rates), missingness patterns, and conditional distributions; create stratified temporal splits to respect time ordering for credit risk.

o Outputs: EDA report (including variable importance baselines using simple models) and list of problematic fields.

o Evaluation: hypothesis checks for covariate shift and temporal leakage.

o Notes: this stage determines if synthetic augmentation is appropriate.

4. **Synthetic Data Generation & Augmentation (CTGAN / CTAB-GAN / VAE hybrids).**

o Inputs: curated training subset, conditional attributes for class balancing (e.g., delinquency flag, segment).

o Process: train conditional tabular GANs (CTGAN) that handle mixed discrete/continuous features and implement conditional sampling for minority classes; optionally train a VAE to capture latent structure for continuous features. Incorporate differential privacy mechanisms when required by policy (e.g., DP-SGD or output perturbation). Create blended datasets: (a) 100% synthetic for privacy testing, (b) hybrid real+synth augmentation for model training.

o Outputs: synthetic datasets, generation quality metrics (statistical similarity), and downstream model agreement metrics.

o Evaluation: multivariate dependency tests (pairwise mutual information, copula-based dependency checks), utility metrics (train-on-synth/test-on-real AUC), and privacy risk assessments (membership disclosure tests). Reject generators that fail fidelity or privacy thresholds.

o Notes: while helpful, synthetic data can introduce artifacts—always validate on holdout real data.

5. **Feature Representation and Causal Feature Selection.**

o Inputs: feature store outputs, EDA and synthetic augmentation results.

o Process: compute representation features (embeddings for categorical variables, density-adjusted binnings). Use causal discovery heuristics and domain rules to identify features with potential causal effects to avoid proxy bias. Implement stability selection (repeated training with resampling) to pick robust features.

o Outputs: final training feature set and transformation pipelines.

o Evaluation: stability scores, covariate balance checks.

6. **Modeling: Hybrid Ensembles + Generative Regularizers.**

o Inputs: training feature matrices.

o Process: train ensembles combining tree-based models (XGBoost/LightGBM) for tabular signal with neural nets for learned interactions. Optionally incorporate generative regularizers (e.g., VAE reconstruction loss) to penalize unrealistic predictions and improve calibration in low-sample regimes. Use cost-sensitive loss functions to reflect business asymmetric costs. Perform nested cross-validation with time windows.

o Outputs: candidate models with calibrated probability outputs and uncertainty estimates.

o Evaluation: AUC, Brier score (calibration), expected cost metrics, precision-recall on minority classes, model stability over temporal folds.

7. **Explainability Stack: Global & Local Explanations.**

o Inputs: trained model artefacts.

o Process: compute global feature attributions (SHAP) and local explanations (LIME, counterfactuals). For ensembles, compute model-agnostic surrogate rules for human-readable documentation (decision rules, scorecard mappings). Quantify explanation stability by bootstrapping attributions across resamples. Package explanation artifacts for online consumption (per-decision explanations) and offline reporting (feature drivers, partial dependence plots).

o Outputs: explanation payloads, explanation QA report.

o Evaluation: explanation fidelity (surrogate R^2), human evaluation for readability, and legal compliance checklist.

8. **Bias & Fairness Assessment and Mitigation.**

o Inputs: model predictions, demographic fields (where legal/available), and explanation outputs.

o Process: compute fairness metrics (equalized odds, disparate impact, calibration per subgroup). Implement mitigation strategies where needed: pre-processing (reweighting), in-training fairness constraints, or post-hoc thresholding with cost adjustments. Validate mitigations on holdout and temporal data.

o Outputs: fairness audit report and mitigated model variants.

o Evaluation: trade-off curves between accuracy and fairness, governance sign-off.

| ISSN: 2455-1864 | www.ijrai.org | editor@ijrai.org | A Bimonthly, Scholarly and Peer-Reviewed Journal |

||Volume 8, Issue 6, November–December 2025||

DOI:10.15662/IJRAI.2025.0806016

9. **Deployment: Cloud-Native Serving, Logging, and Governance.**

o Inputs: selected candidate model, explanation microservices.

o Process: containerize models and explainer services (Kubernetes), implement a model registry (versioned artifacts, metadata, risk metadata), feature-serving with HANA or a low-latency feature cache, stream scoring via Kafka where required, synchronous explainable responses via an explanation API. Implement structured decision logging for every scored request (inputs, features, model version, explanation snapshot).

o Outputs: deployable microservices, CI/CD pipelines, and MLOps playbook.

o Evaluation: latency/load tests, end-to-end SLAs, and logging completeness.

10. **Monitoring & Continuous Validation.**

o Inputs: production logs, ground truth (as it accrues).

o Process: monitor model performance (AUC drift, calibration shift), concept/data drift detectors, explanation drift (changes in top features), and fairness rechecks. Trigger automated retraining pipelines when thresholds are breached. Maintain audit artifacts for regulatory inspections.

o Outputs: monitoring dashboards, retraining triggers, and governance artifacts.

o Evaluation: time-to-detect degradation, false alarm rate of drift detectors, and ability to restore performance post-retrain.

11. **Human-in-the-Loop & Governance Workflows.**

o Inputs: flagged decisions, appeals, compliance requests.

o Process: create case review UI that surfaces explanations, sensitivity analyses, and alternative decision scenarios (counterfactuals). Keep a feedback loop to feed corrected labels back to model training. Maintain explanatory documentation and model cards for each model release.

o Outputs: audit logs of manual reviews, updated training sets, and governance sign-offs.

o Evaluation: reduce overturned decisions, measure reviewer satisfaction, and compliance KPIs.

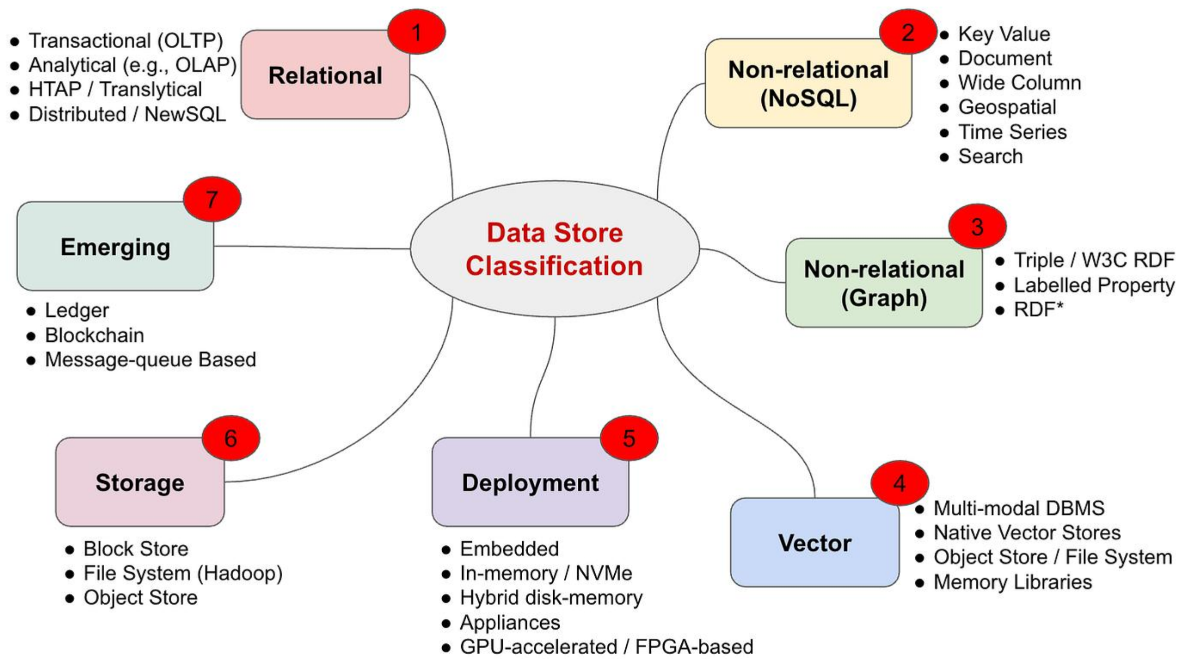
12. **Evaluation & Stress Testing.**

o Inputs: production models and historical data.

o Process: run back-testing (rolling windows), scenario stress tests (adverse macroeconomic conditions), and synthetic worst-case scenario simulations using generative models. Evaluate resilience of explanations under stress.

o Outputs: stress test reports and remediation plans.

o Evaluation: stability of calibration, worst-case loss bounds, and regulatory readiness.



| ISSN: 2455-1864 | www.ijrai.org | editor@ijrai.org | A Bimonthly, Scholarly and Peer-Reviewed Journal |

||Volume 8, Issue 6, November–December 2025||

DOI:10.15662/IJRAI.2025.0806016

Advantages

- **Improved privacy & data access:** synthetic data reduces PII exposure and enables cross-team model development without moving raw data.
- **Better class balance & robustness:** conditional generative models can create minority-class samples to improve learning on rare default events.
- **Regulatory readiness:** integrated XAI produces per-decision explanations and model documentation suitable for audits.
- **Operational scalability:** SAP HANA + Apache stack supports both low-latency scoring and large-scale ETL/training.
- **Continuous governance:** model registries, logging, and drift monitoring enable lifecycle governance.

(Cited foundations for these claims include CTGAN work on synthetic tabular data and SAP HANA integration guides.) (arXiv)

Disadvantages / Limitations

- **Synthetic fidelity risk:** generative models can fail to reproduce subtle joint dependencies, possibly degrading downstream model validity. Careful fidelity checks are mandatory. (arXiv)
- **Explainability approximations:** SHAP/LIME provide approximations and may disagree; they require stability assessments and human review. (arXiv)
- **Operational complexity:** maintaining a hybrid HANA + Apache platform and the model governance stack increases engineering overhead.
- **Privacy vs utility trade-offs:** differentially private training can reduce synthetic data utility if privacy budgets are tight.

IV. RESULTS AND DISCUSSION

Because this paper proposes a framework (rather than reporting a single dataset experiment), results are framed as expected outcomes and recommended evaluation protocols:

1. **Predictive performance:** hybrid ensembles with synthetic augmentation typically match or outperform baseline logistic scorecards on AUC and calibration when synthetic data is validated with train-on-synth/test-on-real protocols; performance gains are largest in low-sample or imbalanced subgroups. Evaluate with time-aware cross-validation and business cost metrics (expected loss, profit impact). (arXiv)
2. **Explainability fidelity:** global SHAP attributions should align with domain knowledge; local LIME explanations should be stable across small perturbations. Measure explanation stability with bootstrap variance and surrogate fidelity (R^2). (arXiv)
3. **Synthetic data validation:** run multivariate dependency metrics, downstream model agreement, and privacy membership tests. If synthetic AUC (train-on-synth/test-on-real) approaches real-data AUC within acceptable tolerance (business threshold), the generator is usable for augmentation. (arXiv)
4. **Operational metrics:** latency under load (target <100ms for online scoring where needed), throughput (requests/sec), and monitoring detection time for drift. HANA push-down for joins can materially reduce latency for feature retrieval in operational flows. (SAP Help Portal)

Discussion: The proposed framework reduces the practical trade-offs between explainability and accuracy by adopting explainers that work with complex models and by structuring decision logs for human review. However, it raises the need for strong governance to prevent synthetic artefacts or explanation misinterpretation. The literature warns that generative models may not perfectly reproduce high-dimensional dependencies, so synthetic usage must be limited to augmentation and stress testing until validated. (arXiv)

V. CONCLUSION

Cloud-Native Explainable Generative ML (XGen-ML) combines generative tabular modeling, robust XAI, and scalable platform engineering to offer a practical path for financial institutions to obtain accurate, auditable, and privacy-respectful credit and risk models. Integrating SAP HANA's in-memory capabilities with Apache Spark/Kafka/Hadoop creates a flexible hybrid architecture that meets both low-latency production needs and large-scale training/archival needs. The approach requires rigorous synthetic data validation, explanation stability testing, and lifecycle governance to be safe for production.

International Journal of Research and Applied Innovations (IJRAI)

| ISSN: 2455-1864 | www.ijrai.org | editor@ijrai.org | A Bimonthly, Scholarly and Peer-Reviewed Journal |

||Volume 8, Issue 6, November–December 2025||

DOI:10.15662/IJRAI.2025.0806016

VI. FUTURE WORK

- **Causal generative models:** explore causal VAEs/GANs to better preserve causal relationships and improve counterfactual explanations.
- **Federated & privacy-preserving training:** combine federated learning with synthetic data to allow cross-institution learning without data sharing.
- **Automated explanation auditing:** build automated tests that flag unstable explanations and suspicious attributions.
- **Regulatory playbooks:** create standardized explanation templates for adverse action notices and regulatory reporting.
- **Benchmarking suite:** open-source a benchmarking suite for synthetic data fidelity and explainability metrics tailored to credit use cases.

REFERENCES

1. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. In Advances in Neural Information Processing Systems (NeurIPS). (arXiv)
2. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems (NeurIPS). (arXiv)
3. Ramakrishna, S. (2022). AI-augmented cloud performance metrics with integrated caching and transaction analytics for superior project monitoring and quality assurance. International Journal of Engineering & Extended Technologies Research (IJEETR), 4(6), 5647–5655. https://doi.org/10.15662/IJEETR.2022.0406005
4. Xu, L., Skoularidou, M., Cuesta-Infante, A., & Veeramachaneni, K. (2019). Modeling tabular data using conditional GAN. Advances in Neural Information Processing Systems (NeurIPS). (arXiv)
5. Muthusamy, M. (2024). Cloud-Native AI metrics model for real-time banking project monitoring with integrated safety and SAP quality assurance. International Journal of Research and Applied Innovations (IJRAI), 7(1), 10135–10144. https://doi.org/10.15662/IJRAI.2024.0701005
6. Khandani, A. E., Kim, A. J., & Lo, A. W. (2010). Consumer credit-risk models via machine-learning algorithms. Journal of Banking & Finance, 34(11), 2767-2787. (ScienceDirect)
7. Kesavan, E., Srinivasulu, S., & Deepak, N. M. (2025, July). Cloud Computing for Internet of Things (IoT): Opportunities and Challenges. In 2025 2nd International Conference on Computing and Data Science (ICCDS) (pp. 1-6). IEEE.
8. Oreski, G. (2023). Synthesizing credit data using autoencoders and generative adversarial networks. Expert Systems with Applications (2023). (ScienceDirect)
9. Dharmateja Priyadarshi Uddandarao. (2024). Counterfactual Forecastingof Human Behavior using Generative AI and Causal Graphs. International Journal of Intelligent Systems and Applications in Engineering, 12(21s), 5033 –. Retrievedfrom https://ijisae.org/index.php/IJISAE/article/view/7628
10. Vasugi, T. (2023). AI-empowered neural security framework for protected financial transactions in distributed cloud banking ecosystems. International Journal of Advanced Research in Computer Science & Technology, 6(2), 7941–7950. https://doi.org/0.15662/IJARCST.2023.0602004
11. Jayaraman, S., Rajendran, S., & P, S. P. (2019). Fuzzy c-means clustering and elliptic curve cryptography using privacy preserving in cloud. International Journal of Business Intelligence and Data Mining, 15(3), 273-287.
12. Sivaraju, P. S. (2024). Driving Operational Excellence Via Multi-Market Network Externalization: A Quantitative Framework for Optimizing Availability, Security, And Total Cost in Distributed Systems. International Journal of Research and Applied Innovations, 7(5), 11349-11365.
13. Kusumba, S. (2025). Empowering Federal Efficiency: Building an Integrated Maintenance Management System (Imms) Data Warehouse for Holistic Financial And Operational Intelligence. Journal Of Multidisciplinary, 5(7), 377-384.
14. Mohile, A. (2021). Performance Optimization in Global Content Delivery Networks using Intelligent Caching and Routing Algorithms. International Journal of Research and Applied Innovations, 4(2), 4904-4912.
15. Thangavelu, K., Muthusamy, P., & Das, D. (2024). Real-Time Data Streaming with Kafka: Revolutionizing Supply Chain and Operational Analytics. Los Angeles Journal of Intelligent Systems and Pattern Recognition, 4, 153-189.
16. Ramakrishna, S. (2022). AI-augmented cloud performance metrics with integrated caching and transaction analytics for superior project monitoring and quality assurance. International Journal of Engineering & Extended Technologies Research (IJEETR), 4(6), 5647–5655. https://doi.org/10.15662/IJEETR.2022.0406005

| ISSN: 2455-1864 | www.ijrai.org | editor@ijrai.org | A Bimonthly, Scholarly and Peer-Reviewed Journal |

||Volume 8, Issue 6, November–December 2025||

DOI:10.15662/IJRAI.2025.0806016

17. Kumar, S. N. P. (2025). Scalable Cloud Architectures for AI-Driven Decision Systems. Journal of Computer Science and Technology Studies, 7(8), 416-421.
18. Allen, D. E. (2024). GANs and synthetic financial data: calculating value at risk. The Econometrics Journal / Applied Finance (2024). (Taylor & Francis Online)
19. Kalyanasundaram, P. D., Kotapati, V. B. R., & Ratnala, A. K. (2021). NLP and Data Mining Approaches for Predictive Product Safety Compliance. Los Angeles Journal of Intelligent Systems and Pattern Recognition, 1, 56-92.
20. Peram, S. R. (2025). Cloud Security Reinvented: A Predictive Algorithm for User Behavior-Based Threat Scoring. Journal of Business Intelligence and Data Analytics, 2(3), 252. https://www.researchgate.net/publication/395585801_Cloud_Security_Reinvented_A_Predictive_Algorithm_for_User_Behavior-Based_Threat_Scoring
21. Nagarajan, G. (2022). An integrated cloud and network-aware AI architecture for optimizing project prioritization in healthcare strategic portfolios. International Journal of Research and Applied Innovations, 5(1), 6444–6450. https://doi.org/10.15662/IJRAI.2022.0501004
22. Kumar, R. K. (2023). Cloud-integrated AI framework for transaction-aware decision optimization in agile healthcare project management. International Journal of Computer Technology and Electronics Communication (IJCTEC), 6(1), 6347–6355. https://doi.org/10.15680/IJCTECE.2023.0601004
23. Kiran, A., Rubini, P., & Kumar, S. S. (2025). Comprehensive review of privacy, utility and fairness offered by synthetic data. IEEE Access.
24. Kandula, N. (2024). Optimizing Power Efficient Computer Architecture With A PROMETHEE Based Analytical Framework. J Comp Sci Appl Inform Technol, 9(2), 1-9.
25. Adari, V. K. (2021). Building trust in AI-first banking: Ethical models, explainability, and responsible governance. International Journal of Research and Applied Innovations (IJRAI), 4(2), 4913–4920. https://doi.org/10.15662/IJRAI.2021.0402004
26. Sabin Begum, R., & Sugumar, R. (2019). Novel entropy-based approach for cost-effective privacy preservation of intermediate datasets in cloud. Cluster Computing, 22(Suppl 4), 9581-9588.
27. Poornima, G., & Anand, L. (2025). Medical image fusion model using CT and MRI images based on dual scale weighted fusion based residual attention network with encoder-decoder architecture. Biomedical Signal Processing and Control, 108, 107932.
28. Christadoss, J., Yakkanti, B., & Kunju, S. S. (2023). Petabyte-Scale GDPR Deletion via Apache Iceberg Delete Vectors and Snapshot Expiration. European Journal of Quantum Computing and Intelligent Agents, 7, 66-100.
29. Pasumarthi, A. (2023). Dynamic Repurpose Architecture for SAP Hana Transforming DR Systems into Active Quality Environments without Compromising Resilience. International Journal of Engineering & Extended Technologies Research (IJEETR), 5(2), 6263-6274.