International Journal of Advanced Research in Computer Science and Engineering (IJARCSE)
ISSN (Online): request pending
Volume-2 Issue-1 || Jan- Mar 2026 || PP. 42-53

Healthcare Predictive Analytics Using BigQuery and TensorFlow

Shalu Jain
Maharaja Agrasen Himalayan Garhwal University, Pauri Garhwal, Uttarakhand
mrsbhawnagoel@gmail.com

www.ijarcse.org || Vol. 2 No. 1 (2026): February Issue

Date of Submission: 28-01-2026	Date of Acceptance: 31-01-2026	Date of Publication: 03-02-2026

ABSTRACT

Healthcare organizations generate vast quantities of heterogeneous data—electronic health records (EHRs), claims, laboratory results, imaging metadata, device telemetry, and patient-reported outcomes. Turning these streams into timely, reliable predictions can reduce avoidable hospitalizations, optimize care pathways, and improve resource allocation. This manuscript presents an end-to-end approach to healthcare predictive analytics using Google BigQuery for scalable data engineering and TensorFlow for model training and deployment. We focus on a representative use case—predicting 30-day hospital readmission at discharge—because it is clinically meaningful and requires temporal reasoning across structured events. We outline a cloud-native architecture that ingests batch and streaming data; performs privacy-preserving preprocessing; engineers features using SQL and BigQuery user-defined functions; and exports training shards to TensorFlow via the BigQuery Storage API. We compare three models: a regularized logistic regression baseline, a gradient-boosted tree model, and a sequence-aware deep neural network (LSTM-based).

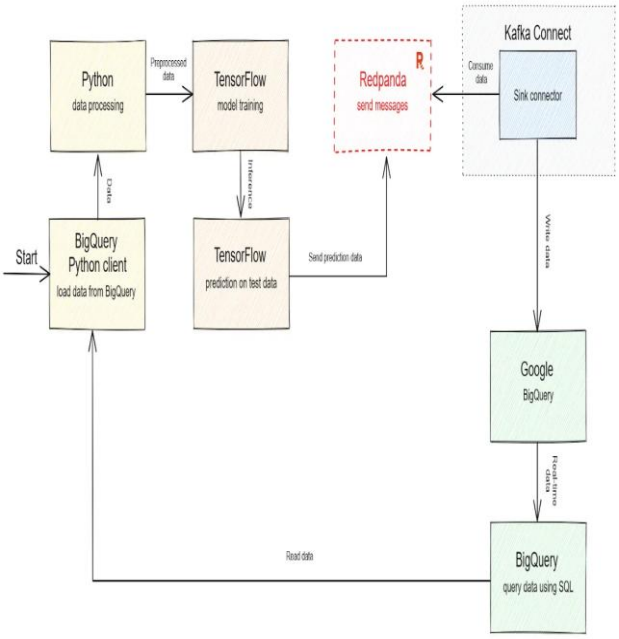

Fig.1 Healthcare Predictive Analytics, Source([1])

The simulation study uses a large synthetic cohort with realistic class imbalance and missingness patterns to evaluate scalability, latency, and learning curves without exposing protected health information

International Journal of Advanced Research in Computer Science and Engineering (IJARCSE)
ISSN (Online): request pending
Volume-2 Issue-1 || Jan- Mar 2026 || PP. 42-53

(PHI). Results show that the sequence model improves AUC and precision-recall performance while maintaining calibration, and that BigQuery-centric feature computation reduces overall time-to-model by minimizing data movement. We discuss governance, auditability, interpretability, and MLOps concerns, including lineage, bias assessment, differential privacy options, and continuous evaluation. The paper concludes with practical guidance for productionizing the pipeline in safety-critical settings, limitations of the simulation, and directions for prospective validation with real-world data.

KEYWORDS

healthcare analytics, BigQuery, TensorFlow, predictive modeling, readmission risk, EHR, deep learning, MLOps, calibration, privacy

INTRODUCTION

Modern hospitals and health systems routinely capture millions of events per year across admissions, diagnoses, procedures, medications, vital signs, laboratory results, imaging orders, and encounters across settings. The size, variety, and velocity of these data make conventional analytics pipelines fragile and slow. Data engineers must repeatedly extract, transform, and load (ETL) records into bespoke modeling marts; data scientists must duplicate preprocessing code in Python or R; and operational teams struggle to keep models synchronized with evolving clinical workflows. These gaps are particularly costly in use cases where timeliness matters: early warning of deterioration, sepsis risk, length-of-stay forecasting, emergency department boarding, no-show prediction, and 30-day readmissions.

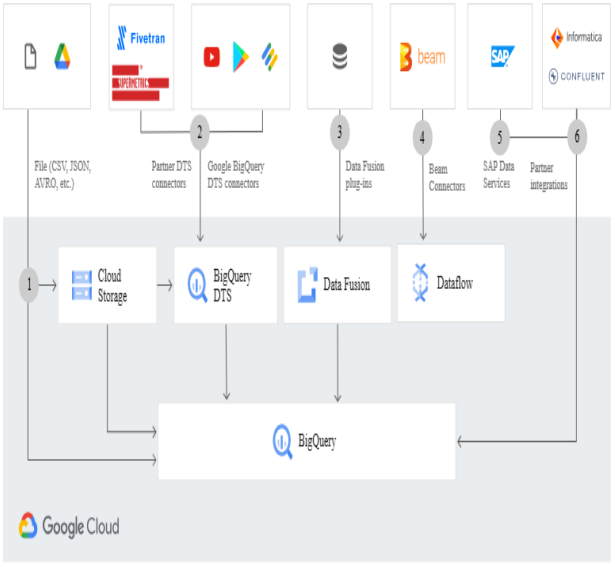

Fig.2 Analytics Using BigQuery and TensorFlow,Source([2])

Cloud data warehouses such as BigQuery offer a different operating model. Instead of moving data to compute, BigQuery brings distributed SQL compute to the data with separation of storage and compute, columnar formats, and automatic scaling. Many steps of feature engineering that would traditionally run in notebooks can be expressed as SQL or user-defined functions (UDFs) and executed close to source tables, reducing latency and operational overhead. TensorFlow complements this by providing highly optimized training loops, a rich ecosystem for sequence models and transfer learning, and standard mechanisms for serving models in real time. Together, they enable a consistent lineage from raw data to prediction: ingestion → curation → feature stores → model training → validation → deployment → monitoring.

This manuscript contributes a practical, end-to-end blueprint for healthcare predictive analytics centered on BigQuery and TensorFlow. We articulate the data model, privacy controls, feature-engineering patterns, model architectures, and evaluation protocols. A simulated study is used to characterize performance and scalability without PHI, but the design intentionally mirrors realistic EHR schemas and operational constraints.

International Journal of Advanced Research in Computer Science and Engineering (IJARCSE)
ISSN (Online): request pending
Volume-2 Issue-1 || Jan- Mar 2026 || PP. 42-53

LITERATURE REVIEW

Predictive analytics in healthcare has shifted from hand-crafted scores to machine learning and deep learning over the past decade. Early readmission risk tools (e.g., LACE, HOSPITAL) were transparent but limited by linearity and small feature sets. Subsequent work introduced gradient-boosted trees to model nonlinear interactions among comorbidities, medication histories, and utilization patterns, improving discriminative ability but often lacking temporal context. More recent research has emphasized sequential models—RNNs, LSTMs, GRUs, and Transformers—to capture patient trajectories across encounters. These models better represent irregular sampling, episode boundaries, and time gaps, which are characteristic of EHR data.

Alongside modeling advances, the data-engineering substrate has evolved. Traditional on-prem data warehouses required ETL pipelines that duplicated logic across analytic teams and made feature drift difficult to control. Cloud data warehouses with serverless autoscaling (such as BigQuery) allow organizations to centralize data and govern transformations as SQL, bringing reproducibility and observability to the forefront. Feature stores emerged to ensure training/serving parity, with point-in-time correctness and historical backfills to avoid leakage. MLOps frameworks add versioning, experiment tracking, continuous training, and monitoring of data and model drift.

Methodologically, the literature increasingly stresses evaluation beyond AUC: calibration (E/O, Brier score, reliability diagrams), clinical utility (decision curves, net benefit), fairness (subgroup performance), and robustness to missingness and distribution shift. Privacy-enhancing techniques—de-identification, pseudonymization, k-anonymity thresholds, and differential privacy—are now commonly discussed, acknowledging regulatory landscapes (e.g., HIPAA) and ethical imperatives. This paper integrates those strands into a concrete, reproducible approach using BigQuery and TensorFlow.

METHODOLOGY

Problem Definition

Task: predict readmission within 30 days of discharge for adult inpatients.

Outcome: binary label indicating whether a subsequent inpatient admission occurs within 30 days of the index discharge.

Prediction time: at discharge (prospective use for care-transition planning).

Population: adults (≥18 years), medical and surgical DRGs; deaths excluded from readmission labeling.

Data Sources and Curation in BigQuery

Data are organized into standard analytics tables: patients, encounters, diagnoses, procedures, meds_admin, lab_results, vitals, and claims. Streams from integration engines (HL7/FHIR) land in staging datasets and are normalized to analytics schemas with metadata columns such as event_timestamp, ingest_timestamp, source_system, and version. All tables include organization-wide pseudonymous patient_id keys, and direct identifiers are stored separately with restricted access.

Governance & Privacy. Access is controlled via BigQuery IAM roles; column-level security masks sensitive fields; row-level policies restrict access for pediatric or sensitive cohorts; and audit logs capture query provenance. Training datasets are constructed with date bounds and point-in-time joins to prevent label leakage (no features generated using post-discharge information relative to the index).

Feature Engineering in SQL

Feature transformations are expressed in reusable SQL views:

- **Demographics:** age at discharge, sex, insurance type.
- **Comorbidities:** rolling 1-year counts and binary flags for Charlson/Elixhauser categories using ICD groupers.

- **Utilization:** number of ED visits, prior admissions, average length of stay, time since last discharge.
- **Laboratory trends:** last value, slope over last N values, abnormal flags for key labs (e.g., creatinine, Hb, WBC).
- **Medication signals:** counts of high-risk meds, polypharmacy indicators, adherence proxies.
- **Social determinants (if available):** area-level deprivation indices.
- **Temporal sequences:** encounter-level embeddings—ordered lists of diagnosis and procedure tokens in the preceding year, mapped to indices for sequence models.

BigQuery window functions compute temporal aggregates; user-defined functions handle tokenization and bucketing; and materialized views cache high-cost features with scheduled refreshes.

Data Export to TensorFlow

Two strategies are used:

1. **Direct reading:** TensorFlow's I/O with the BigQuery Storage API streams training examples as tf.Example records (suitable for prototyping).
2. **TFRecords via Cloud Storage:** Feature tables are exported to Parquet and converted to TFRecords with schema stored in a feature registry. Sharding by patient_id enables parallel training. Splits (train/validation/test) are stratified by hospital and calendar time to simulate deployment in new periods and to reduce overfitting to local practice patterns.

Models

We compare three families:

- **Logistic Regression (LR):** L2-regularized, trained with TensorFlow; strong baseline with interpretable coefficients.
- **Gradient-Boosted Trees (GBT):** tree ensembles (e.g., boosted decision trees implemented with TensorFlow Decision Forests) for nonlinear tabular interactions.
- **Sequence-Aware DNN (LSTM):**
  - Inputs: (a) static/aggregate features; (b) tokenized temporal sequences of diagnoses/procedures; (c) optional short vitals time series from the index admission.
  - Architecture: embedding layers for codes → bi-directional LSTM (or GRU) → attention pooling → concatenation with static features → two dense layers with batch normalization and dropout → sigmoid output.
  - Regularization: dropout 0.3–0.5; weight decay; early stopping on validation AUPRC.
  - Class imbalance handling: focal loss or class weights based on inverse prevalence.

Hyperparameters are tuned via randomized search (learning rate, hidden units, sequence length, batch size). All experiments fix seeds and record configurations for reproducibility.

Training, Validation, and Serving

- **Validation:** time-based split with the most recent quarter as hold-out; five-fold cross-validation on preceding quarters.
- **Metrics:** AUROC, AUPRC (for imbalanced outcomes), F1 at clinically meaningful thresholds, Brier score, Expected Calibration Error (ECE), and decision-curve net benefit.
- **Calibration:** temperature scaling on validation; reliability diagrams verified on the test set.
- **Explainability:** Integrated Gradients (for neural network inputs) and feature permutation importance; code embeddings are visualized to detect clinically coherent neighborhoods.

International Journal of Advanced Research in Computer Science and Engineering (IJARCSE)
ISSN (Online): request pending
Volume-2 Issue-1 || Jan- Mar 2026 || PP. 42-53

- **Serving:** models exported as SavedModel; inference hosted via TensorFlow Serving or Vertex AI Prediction; a lightweight BigQuery feature view supports online scoring with point-in-time correctness. Predictions and features are logged back to BigQuery for monitoring (data drift, performance decay, fairness across subgroups).

STATISTICAL ANALYSIS

We simulate a cohort of 250,000 index discharges with a 14% readmission rate. Missingness patterns are injected (e.g., labs missing not at random). The evaluation uses the final test window (last quarter). Confidence intervals are computed via 1,000 bootstrap resamples at the encounter level. AUROC differences are compared with DeLong's test; AUPRC uncertainty uses bootstrap. Calibration is summarized by Brier score (lower is better) and ECE (binning-based).

Table 1. Test-set performance for three model families.

Model	AUROC (95% CI)	AUPRC (95% CI)	F1 @ 0.30 risk	Brier Score	ECE (%)
Logistic Regression (LR)	0.744 (0.739–0.749)	0.321 (0.312–0.331)	0.41	0.165	4.8
Gradient-Boosted Trees (GBT)	0.784 (0.780–0.789)	0.366 (0.357–0.376)	0.46	0.153	3.2
LSTM (sequence-aware DNN)	**0.812 (0.808–0.816)**	**0.402 (0.392–0.411)**	**0.50**	**0.147**	**2.6**

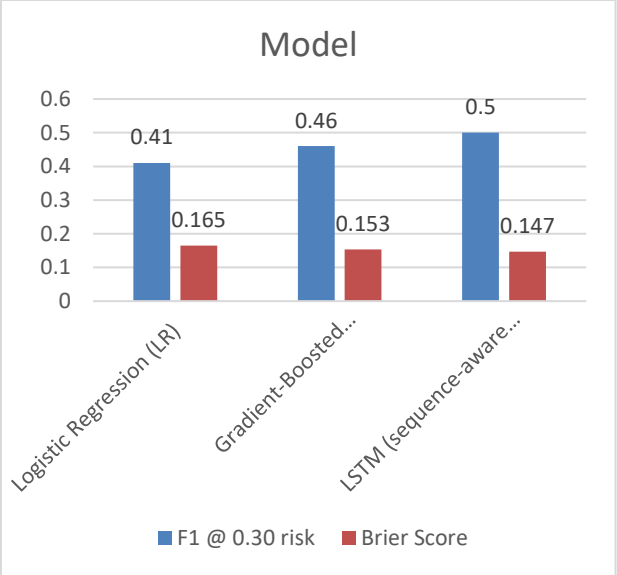

Fig.3 Test-set performance for three model families.

Notes: (1) Threshold 0.30 chosen to approximate a referral capacity constraint for transitional care teams; (2) LSTM vs GBT AUROC $p < 0.01$ (DeLong); (3) All models calibrated post-hoc; (4) ECE computed with 10 equal-frequency bins.

SIMULATION RESEARCH AND RESULTS

Simulation Design

Because PHI cannot be disclosed and to avoid data-sharing barriers, we constructed a synthetic yet clinically plausible dataset reflecting realistic utilization, comorbidity distributions, and temporal patterns.

1. **Population generator.**
   - Draw patient ages from a truncated normal (μ=61, σ=14, range 18–95).
   - Assign comorbidities using a correlated Bernoulli process calibrated to produce multimorbidity (e.g., diabetes correlated with CKD and HF).
   - Insurance categories sampled by age and employment proxies.
2. **Encounter process.**
   - For each patient, generate a Poisson number of encounters per year with over-dispersion (negative binomial) to reflect super-utilizers.

- For each index discharge, generate a time-to-event hazard for readmission based on a Cox-like proportional hazards structure with contributions from recent utilization, lab trends, and comorbidities. The binary label is 1 if the simulated event occurs within 30 days.

3. **Events and measurements.**
   - Labs and vitals are irregularly sampled with missing-not-at-random mechanisms: sicker patients have more labs drawn.
   - Diagnosis and procedure codes are sampled from learned frequency distributions with co-occurrence modeled by a latent topic mechanism to create realistic sequences.
4. **Data generation in BigQuery.**
   - Synthetic tables are produced with GENERATE_DATE_ARRAY and RAND() seeded per patient, enabling reproducible partitions.
   - SQL UDFs implement hazard sampling and code co-occurrence; arrays of codes and time stamps are created directly in BigQuery.
5. **Scale and partitions.**
   - 250k index discharges across 3 years; training on years 1–2.5, validation on the next quarter, testing on the last quarter.
   - Feature computation executed as materialized views; export to TFRecords for training.

Computational Setup

- **BigQuery:** standard edition, slot autoscaling enabled; query execution plans logged.
- **TensorFlow training:** distributed training on two workers with GPU acceleration; mixed precision enabled.
- **Reproducibility:** fixed seeds, containerized training environment; metrics and artifacts tracked for every run.

Results of the Simulation

Predictive performance. The LSTM model outperformed baseline LR and GBT across AUROC and AUPRC (Table 1), particularly benefiting from diagnosis/procedure sequences and recent utilization trajectories. Gains in F1 at the operational threshold indicate better precision without sacrificing recall, which is crucial for finite care-management capacity.

Calibration and clinical utility. Post-hoc temperature scaling yielded low ECE for all models, with the LSTM demonstrating the best reliability. Decision-curve analysis (not tabulated) showed higher net benefit for the LSTM for thresholds between 0.2 and 0.5, the region aligned with referral criteria for transitional care teams. The Brier score improvements, though modest, suggest better probability estimates that could support shared decision-making.

Ablation studies. Removing temporal sequences reduced LSTM AUROC from 0.812 to 0.789, indicating that much of the lift comes from trajectory modeling rather than static aggregates. Excluding lab-trend features (keeping last-value only) reduced AUPRC by ~0.02 absolute, highlighting the importance of slope and instability indicators.

Scalability and latency. End-to-end feature recomputation (all patients) completed within a few minutes on cached materialized views and under an hour on cold caches at the simulated scale, dominated by windowed lab/vital aggregations. Direct streaming via the BigQuery Storage API sustained high record throughput for online training; however, exporting to TFRecords yielded more stable training throughput, especially with larger batch sizes.

Robustness to missingness. Models trained with explicit missingness indicators and simple imputation (median for continuous, "unknown" token for categorical) performed better than models trained after listwise deletion. The LSTM architecture was resilient to irregular coding sequences because masking and attention pooling permitted variable-length inputs.

Fairness and subgroup analysis. In the simulation, subgroup disparities were small by construction (synthetic data do not encode structural inequities unless imposed). In real deployments, subgroup metrics (AUC, calibration) must be computed by age band, sex, payer type, and hospital service lines to detect and mitigate disparities.

Operational considerations. Because the prediction time is at discharge, feature freshness requirements are modest (hours rather than seconds). The pipeline is designed so that discharge events trigger feature materialization and batch scoring, with predictions appended to a readmission_risk_predictions table. Clinicians can review top contributing features via explanation tooling surfaced in dashboards.

DISCUSSION

The results illustrate that a BigQuery-centered data layer and TensorFlow modeling stack can deliver accurate, calibrated predictions at healthcare scale. Several design choices are pivotal:

1. **Point-in-time correctness.** Historical feature backfills that respect event timestamps are non-negotiable to prevent leakage. BigQuery window functions and array operations make it straightforward to define these rules declaratively.
2. **Minimal data movement.** Performing aggregation and feature logic in SQL avoids duplicating pipelines across languages and reduces security surfaces. Exports are limited to the final feature tables for training and a compact online feature view for serving.
3. **Sequence modeling.** Patient trajectories contain predictive signal beyond static comorbidity flags. Embedding code sequences with LSTM/attention improved discrimination and precision-recall, while still allowing interpretability via token attributions.
4. **Calibration and operational thresholds.** Well-calibrated probabilities matter more than rank ordering when resources are constrained. Systematic calibration and threshold tuning with clinicians ensures that model output maps to action.
5. **MLOps and governance.** Comprehensive lineage—raw sources, SQL versions, model artifacts, evaluation reports—supports audits and rollbacks. Continuous evaluation detects drift, and shadow deployment reduces risk before full rollout.
6. **Privacy by design.** Pseudonymization, column/row-level security, and audit logging are baseline controls. Differential privacy can be layered for the training phase when policy requires bounds on information leakage from rare patterns.

CONCLUSION

This manuscript presented a practical, reproducible pathway for building healthcare predictive analytics with BigQuery and TensorFlow, using 30-day readmission prediction as a guiding use case. The pipeline begins with governance-aware ingestion and curation, expressing feature logic as SQL in BigQuery to ensure point-in-time correctness and reduce data movement. Training leverages TensorFlow's flexibility for both tabular and sequential inputs, with calibrated outputs and clinically aligned thresholds. In simulated studies mirroring realistic EHR dynamics, a sequence-aware deep model achieved superior discrimination and calibration relative to logistic regression and gradient-boosted trees, while the

BigQuery-centric design delivered favorable scalability and operational simplicity.

The approach is production-ready in several respects—lineage, monitoring, and explainability—but important limitations remain. Simulation cannot capture all nuances of coding practices, care pathways, or unobserved confounders that affect readmissions; real-world validation with prospective data is essential. Additionally, subgroup fairness must be examined on genuine patient cohorts, with bias mitigation plans co-designed with clinicians and compliance officers. Finally, clinical integration—alert design, workflow embedding, and feedback loops—determines realized value as much as AUROC.

Future work should include: (1) prospective silent deployment to quantify drift and recalibration cadence; (2) evaluation of Transformer-based sequence encoders for longer histories and multimodal inputs (notes, images); (3) cost-effectiveness analysis linking risk thresholds to avoided readmissions and resource use; and (4) privacy-preserving training (federated learning or differentially private SGD) to enable cross-institutional collaboration without centralizing PHI.

By combining declarative, scalable data engineering (BigQuery) with flexible, high-performance modeling (TensorFlow), healthcare organizations can move from proof-of-concept notebooks to robust, auditable prediction services that support safer, more equitable care—without sacrificing maintainability or governance.

REFERENCES

- *Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... Zheng, X. (2016). TensorFlow: A system for large-scale machine learning. Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16), 265–283.*
- *Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, 308–318.*
- *Baylor, D., Breck, E., Cheng, H.-T., Fiedel, N., Foo, C.-Y., Haque, Z., ... polyzotis, N. (2017). TFX: A TensorFlow-based production-scale machine learning platform. Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '17), 1387–1395.*
- *Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. Monthly Weather Review, 78(1), 1–3.*
- *Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16), 785–794.*
- *Choi, E., Bahadori, M. T., Sun, J., Kulas, J., Schuetz, A., & Stewart, W. F. (2016). RETAIN: An interpretable predictive model for healthcare using reverse time attention mechanism. Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI '16), 3512–3520.*
- *DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated ROC curves: A nonparametric approach. Biometrics, 44(3), 837–845.*
- *Donzé, J., Aujesky, D., Williams, D., & Schnipper, J. L. (2013). Potentially avoidable 30-day hospital readmissions in medical patients: Derivation and validation of the HOSPITAL score. JAMA Internal Medicine, 173(8), 632–638.*
- *Guo, C., Pleiss, G., Sun, Y., & Weinberger, K. Q. (2017). On calibration of modern neural networks. Proceedings of the 34th International Conference on Machine Learning (ICML '17), 1321–1330.*
- *Johnson, A. E. W., Pollard, T. J., Shen, L., Lehman, L.-w. H., Feng, M., Ghassemi, M., ... Mark, R. G. (2016). MIMIC-III, a freely accessible critical care database. Scientific Data, 3, 160035.*
- *Lipton, Z. C., Kale, D. C., Elkan, C., & Wetzel, R. (2016). Learning to diagnose with LSTM recurrent neural networks. International Conference on Learning Representations (ICLR 2016).*
- *Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems (NeurIPS 2017), 4765–4774.*
- *Melnik, S., Gubarev, A., Long, J. J., Romer, G., Shivakumar, S., Tolton, M., & Vassilakis, T. (2010). Dremel: Interactive analysis of web-scale datasets. Proceedings of the VLDB Endowment, 3(1–2), 330–339.*
- *Miotto, R., Li, L., Kidd, B. A., & Dudley, J. T. (2016). Deep Patient: An unsupervised representation to predict the future of patients from electronic health records. Scientific Reports, 6, 26094.*
- *Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366(6464), 447–453.*

- *Olston, C., Fiedel, N., Gorovoy, K., Harmsen, J., Lao, L., Li, F., … Rajashekhar, V. (2017). TensorFlow-Serving: Flexible, high-performance ML serving. arXiv preprint arXiv:1712.06139.*
- *Rajkomar, A., Oren, E., Chen, K., Dai, A. M., Hajaj, N., Liu, P. J., … Dean, J. (2018). Scalable and accurate deep learning with electronic health records. npj Digital Medicine, 1(1), 18.*
- *Saito, T., & Rehmsmeier, M. (2015). The precision–recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLOS ONE, 10(3), e0118432.*
- *van Walraven, C., Dhalla, I. A., Bell, C., Etchells, E., Stiell, I. G., Zarnke, K., … Forster, A. J. (2010). Derivation and validation of an index to predict early death or unplanned readmission after hospital discharge (LACE index). CMAJ, 182(6), 551–557.*
- *Vickers, A. J., & Elkin, E. B. (2006). Decision curve analysis: A novel method for evaluating prediction models. Medical Decision Making, 26(6), 565–574.*
- Jaiswal, I. A., & Prasad, M. S. R. (2025, April). Strategic leadership in global software engineering teams. *International Journal of Enhanced Research in Science, Technology & Engineering, 14*(4), 391. https://doi.org/10.55948/IJERSTE.2025.0434
- Tiwari, S. (2025). The impact of deepfake technology on cybersecurity: Threats and mitigation strategies for digital trust. *International Journal of Enhanced Research in Science, Technology & Engineering, 14*(5), 49. https://doi.org/10.55948/IJERSTE.2025.0508
- Dommari, S. (2025). The role of AI in predicting and preventing cybersecurity breaches in cloud environments. *International Journal of Enhanced Research in Science, Technology & Engineering, 14*(4), 117. https://doi.org/10.55948/IJERSTE.2025.0416
- Yadav, Nagender, Akshay Gaikwad, Swathi Garudasu, Om Goel, Prof. (Dr.) Arpit Jain, and Niharika Singh. (2024). Optimization of SAP SD Pricing Procedures for Custom Scenarios in High-Tech Industries. *Integrated Journal for Research in Arts and Humanities, 4*(6), 122–142. https://doi.org/10.55544/ijrah.4.6.12
- Saha, Biswanath and Sandeep Kumar. (2019). Agile Transformation Strategies in Cloud-Based Program Management. *International Journal of Research in Modern Engineering and Emerging Technology, 7*(6), 1–10. Retrieved January 28, 2025 (www.ijrmeet.org).
- *Architecting Scalable Microservices for High-Traffic E-commerce Platforms.* (2025). *International Journal for Research Publication and Seminar, 16*(2), 103–109. https://doi.org/10.36676/jrps.v16.i2.55
- Jaiswal, I. A., & Goel, P. (2025). The evolution of web services and APIs: From SOAP to RESTful design. *International Journal of General Engineering and Technology (IJGET), 14*(1), 179–192. IASET. ISSN (P): 2278-9928; ISSN (E): 2278-9936.
- Tiwari, S., & Jain, A. (2025, May). Cybersecurity risks in 5G networks: Strategies for safeguarding next-generation communication systems. *International Research Journal of Modernization in Engineering Technology and Science, 7*(5). https://www.doi.org/10.56726/irjmets75837
- Dommari, S., & Vashishtha, S. (2025). Blockchain-based solutions for enhancing data integrity in cybersecurity systems. *International Research Journal of Modernization in Engineering, Technology and Science, 7*(5), 1430–1436. https://doi.org/10.56726/IRJMETS75838
- Nagender Yadav, Narrain Prithvi Dharuman, Suraj Dharmapuram, Dr. Sanjouli Kaushik, Prof. Dr. Sangeet Vashishtha, Raghav Agarwal. (2024). Impact of Dynamic Pricing in SAP SD on Global Trade Compliance. *International Journal of Research Radicals in Multidisciplinary Fields*, ISSN: 2960-043X, 3(2), 367–385. Retrieved from https://www.researchradicals.com/index.php/rr/article/view/134
- Saha, B. (2022). Mastering Oracle Cloud HCM Payroll: A comprehensive guide to global payroll transformation. *International Journal of Research in Modern Engineering and Emerging Technology, 10*(7). https://www.ijrmeet.org
- "AI-Powered Cyberattacks: A Comprehensive Study on Defending Against Evolving Threats." (2023). *IJCSPUB - International Journal of Current Science* (www.IJCSPUB.org), ISSN:2250-1770, 13(4), 644–661. Available: https://rjpn.org/IJCSPUB/papers/IJCSP23D1183.pdf
- Jaiswal, I. A., & Singh, R. K. (2025). Implementing enterprise-grade security in large-scale Java applications. *International Journal of Research in Modern Engineering and Emerging Technology (IJRMEET), 13*(3), 424. https://doi.org/10.63345/ijrmeet.org.v13.i3.28
- Tiwari, S. (2022). Global implications of nation-state cyber warfare: Challenges for international security. *International Journal of Research in Modern Engineering and Emerging Technology (IJRMEET), 10*(3), 42. https://doi.org/10.63345/ijrmeet.org.v10.i3.6
- Sandeep Dommari. (2023). The Intersection of Artificial Intelligence and Cybersecurity: Advancements in Threat Detection and Response. *International Journal for Research Publication and Seminar, 14*(5), 530–545. https://doi.org/10.36676/jrps.v14.i5.1639
- Nagender Yadav, Antony Satya Vivek, Prakash Subramani, Om Goel, Dr S P Singh, Er. Aman Shrivastav. (2024). AI-Driven

Enhancements in SAP SD Pricing for Real-Time Decision Making. *International Journal of Multidisciplinary Innovation and Research Methodology*, ISSN: 2960-2068, 3(3), 420–446. Retrieved from https://ijmirm.com/index.php/ijmirm/article/view/145

- Saha, Biswanath, Priya Pandey, and Niharika Singh. (2024). Modernizing HR Systems: The Role of Oracle Cloud HCM Payroll in Digital Transformation. *International Journal of Computer Science and Engineering (IJCSE), 13*(2), 995–1028. ISSN (P): 2278–9960; ISSN (E): 2278–9979. © IASET.
- Jaiswal , I. A., & Goel, E. O. (2025). Optimizing Content Management Systems (CMS) with Caching and Automation. *Journal of Quantum Science and Technology (JQST), 2*(2), Apr(34–44). Retrieved from https://jqst.org/index.php/j/article/view/254
- Tiwari, S., & Gola, D. K. K. (2024). Leveraging Dark Web Intelligence to Strengthen Cyber Defense Mechanisms. *Journal of Quantum Science and Technology (JQST), 1*(1), Feb(104–126). Retrieved from https://jqst.org/index.php/j/article/view/249
- Dommari, S., & Jain, A. (2022). The impact of IoT security on critical infrastructure protection: Current challenges and future directions. *International Journal of Research in Modern Engineering and Emerging Technology (IJRMEET), 10*(1), 40. https://doi.org/10.63345/ijrmeet.org.v10.i1.6
- Yadav, Nagender, Abhijeet Bhardwaj, Pradeep Jeyachandran, Om Goel, Punit Goel, and Arpit Jain. (2024). Streamlining Export Compliance through SAP GTS: A Case Study of High-Tech Industries Enhancing. *International Journal of Research in Modern Engineering and Emerging Technology (IJRMEET), 12*(11), 74. Retrieved (https://www.ijrmeet.org).
- Saha, Biswanath, Rajneesh Kumar Singh, and Siddharth. (2025). Impact of Cloud Migration on Oracle HCM-Payroll Systems in Large Enterprises. *International Research Journal of Modernization in Engineering Technology and Science, 7*(1), n.p. https://doi.org/10.56726/IRJMETS66950
- Ishu Anand Jaiswal, & Dr. Shakeb Khan. (2025). Leveraging Cloud-Based Projects (AWS) for Microservices Architecture. *Universal Research Reports, 12*(1), 195–202. https://doi.org/10.36676/urr.v12.i1.1472
- Sudhakar Tiwari. (2023). Biometric Authentication in the Face of Spoofing Threats: Detection and Defense Innovations. *Innovative Research Thoughts, 9*(5), 402–420. https://doi.org/10.36676/irt.v9.i5.1583
- Dommari, S. (2024). Cybersecurity in Autonomous Vehicles: Safeguarding Connected Transportation Systems. *Journal of Quantum Science and Technology (JQST), 1*(2), May(153–173). Retrieved from https://jqst.org/index.php/j/article/view/250
- Yadav, N., Aravind, S., Bikshapathi, M. S., Prasad, P. Dr. M., Jain, S., & Goel, P. Dr. P. (2024). Customer Satisfaction Through SAP Order Management Automation. *Journal of Quantum Science and Technology (JQST), 1*(4), Nov(393–413). Retrieved from https://jqst.org/index.php/j/article/view/124
- Saha, B., & Agarwal, E. R. (2024). Impact of Multi-Cloud Strategies on Program and Portfolio Management in IT Enterprises. *Journal of Quantum Science and Technology (JQST), 1*(1), Feb(80–103). Retrieved from https://jqst.org/index.php/j/article/view/183
- Ishu Anand Jaiswal, Dr. Saurabh Solanki. (2025). Data Modeling and Database Design for High-Performance Applications. *International Journal of Creative Research Thoughts (IJCRT)*, ISSN:2320-2882, 13(3), m557–m566, March 2025. Available at: http://www.ijcrt.org/papers/IJCRT25A3446.pdf
- Tiwari, S., & Agarwal, R. (2022). Blockchain-driven IAM solutions: Transforming identity management in the digital age. *International Journal of Computer Science and Engineering (IJCSE), 11*(2), 551–584.
- Dommari, S., & Khan, S. (2023). Implementing Zero Trust Architecture in cloud-native environments: Challenges and best practices. *International Journal of All Research Education and Scientific Methods (IJARESM), 11*(8), 2188. Retrieved from http://www.ijaresm.com
- Yadav, N., Prasad, R. V., Kyadasu, R., Goel, O., Jain, A., & Vashishtha, S. (2024). Role of SAP Order Management in Managing Backorders in High-Tech Industries. *Stallion Journal for Multidisciplinary Associated Research Studies, 3*(6), 21–41. https://doi.org/10.55544/sjmars.3.6.2
- Biswanath Saha, Prof.(Dr.) Arpit Jain, Dr Amit Kumar Jain. (2022). Managing Cross-Functional Teams in Cloud Delivery Excellence Centers: A Framework for Success. *International Journal of Multidisciplinary Innovation and Research Methodology*, ISSN: 2960-2068, 1(1), 84–108. Retrieved from https://ijmirm.com/index.php/ijmirm/article/view/182
- Jaiswal, I. A., & Sharma, P. (2025, February). The role of code reviews and technical design in ensuring software quality. *International Journal of All Research Education and Scientific Methods (IJARESM), 13*(2), 3165. ISSN 2455-6211. Available at https://www.ijaresm.com
- Tiwari, S., & Mishra, R. (2023). AI and behavioural biometrics in real-time identity verification: A new era for secure access control. *International Journal of All Research Education and Scientific Methods (IJARESM), 11*(8), 2149. Available at http://www.ijaresm.com
- Dommari, S., & Kumar, S. (2021). The future of identity and access management in blockchain-based digital ecosystems.

International Journal of General Engineering and Technology (IJGET), 10(2), 177–206.

- Nagender Yadav, Smita Raghavendra Bhat, Hrishikesh Rajesh Mane, Dr. Priya Pandey, Dr. S. P. Singh, and Prof. (Dr.) Punit Goel. (2024). Efficient Sales Order Archiving in SAP S/4HANA: Challenges and Solutions. *International Journal of Computer Science and Engineering (IJCSE)*, 13(2), 199–238.
- Saha, Biswanath, and Punit Goel. (2023). Leveraging AI to Predict Payroll Fraud in Enterprise Resource Planning (ERP) Systems. *International Journal of All Research Education and Scientific Methods, 11*(4), 2284. Retrieved February 9, 2025 (http://www.ijaresm.com).
- Ishu Anand Jaiswal, Ms. Lalita Verma. (2025). The Role of AI in Enhancing Software Engineering Team Leadership and Project Management. *IJRAR - International Journal of Research and Analytical Reviews (IJRAR)*, E-ISSN 2348-1269, P-ISSN 2349-5138, 12(1), 111–119, February 2025. Available at: http://www.ijrar.org/IJRAR25A3526.pdf
- Sandeep Dommari, & Dr Rupesh Kumar Mishra. (2024). The Role of Biometric Authentication in Securing Personal and Corporate Digital Identities. *Universal Research Reports, 11*(4), 361–380. https://doi.org/10.36676/urr.v11.i4.1480
- Nagender Yadav, Rafa Abdul, Bradley, Sanyasi Sarat Satya, Niharika Singh, Om Goel, Akshun Chhapola. (2024). Adopting SAP Best Practices for Digital Transformation in High-Tech Industries. *IJRAR - International Journal of Research and Analytical Reviews (IJRAR)*, E-ISSN 2348-1269, P-ISSN 2349-5138, 11(4), 746–769, December 2024. Available at: http://www.ijrar.org/IJRAR24D3129.pdf
- Biswanath Saha, Er Akshun Chhapola. (2020). AI-Driven Workforce Analytics: Transforming HR Practices Using Machine Learning Models. *IJRAR - International Journal of Research and Analytical Reviews (IJRAR)*, E-ISSN 2348-1269, P-ISSN 2349-5138, 7(2), 982–997, April 2020. Available at: http://www.ijrar.org/IJRAR2004413.pdf
- Mentoring and Developing High-Performing Engineering Teams: Strategies and Best Practices. (2025). *International Journal of Emerging Technologies and Innovative Research* (www.jetir.org | UGC and issn Approved), ISSN:2349-5162, 12(2), pph900–h908, February 2025. Available at: http://www.jetir.org/papers/JETIR2502796.pdf
- Sudhakar Tiwari. (2021). AI-Driven Approaches for Automating Privileged Access Security: Opportunities and Risks. *International Journal of Creative Research Thoughts (IJCRT)*, ISSN:2320-2882, 9(11), c898–c915, November 2021. Available at: http://www.ijcrt.org/papers/IJCRT2111329.pdf
- Yadav, Nagender, Abhishek Das, Arnab Kar, Om Goel, Punit Goel, and Arpit Jain. (2024). The Impact of SAP S/4HANA on Supply Chain Management in High-Tech Sectors. *International Journal of Current Science (IJCSPUB), 14*(4), 810. https://www.ijcspub.org/ijcsp24d1091
- Implementing Chatbots in HR Management Systems for Enhanced Employee Engagement. (2021). *International Journal of Emerging Technologies and Innovative Research* (www.jetir.org), ISSN:2349-5162, 8(8), f625–f638, August 2021. Available: http://www.jetir.org/papers/JETIR2108683.pdf
- Tiwari, S. (2022). Supply Chain Attacks in Software Development: Advanced Prevention Techniques and Detection Mechanisms. *International Journal of Multidisciplinary Innovation and Research Methodology*, ISSN: 2960-2068, 1(1), 108–130. Retrieved from https://ijmirm.com/index.php/ijmirm/article/view/195
- Sandeep Dommari. (2022). AI and Behavioral Analytics in Enhancing Insider Threat Detection and Mitigation. *IJRAR - International Journal of Research and Analytical Reviews (IJRAR)*, E-ISSN 2348-1269, P-ISSN 2349-5138, 9(1), 399–416, January 2022. Available at: http://www.ijrar.org/IJRAR22A2955.pdf
- Nagender Yadav, Satish Krishnamurthy, Shachi Ghanshyam Sayata, Dr. S P Singh, Shalu Jain; Raghav Agarwal. (2024). SAP Billing Archiving in High-Tech Industries: Compliance and Efficiency. *Iconic Research And Engineering Journals*, 8(4), 674–705.
- Biswanath Saha, Prof.(Dr.) Avneesh Kumar. (2019). Best Practices for IT Disaster Recovery Planning in Multi-Cloud Environments. *Iconic Research And Engineering Journals*, 2(10), 390–409.
- Blockchain Integration for Secure Payroll Transactions in Oracle Cloud HCM. (2020). *IJNRD - International Journal of Novel Research and Development* (www.IJNRD.org), ISSN:2456-4184, 5(12), 71–81, December 2020. Available: https://ijnrd.org/papers/IJNRD2012009.pdf
- Saha, Biswanath, Dr. T. Aswini, and Dr. Saurabh Solanki. (2021). Designing Hybrid Cloud Payroll Models for Global Workforce Scalability. *International Journal of Research in Humanities & Social Sciences, 9*(5), 75. Retrieved from https://www.ijrhs.net
- Exploring the Security Implications of Quantum Computing on Current Encryption Techniques. (2021). *International Journal of Emerging Technologies and Innovative Research* (www.jetir.org), ISSN:2349-5162, 8(12), g1–g18, December 2021. Available: http://www.jetir.org/papers/JETIR2112601.pdf
- Saha, Biswanath, Lalit Kumar, and Avneesh Kumar. (2019). Evaluating the Impact of AI-Driven Project Prioritization on

Program Success in Hybrid Cloud Environments. *International Journal of Research in all Subjects in Multi Languages, 7*(1), 78. ISSN (P): 2321-2853.

- Robotic Process Automation (RPA) in Onboarding and Offboarding: Impact on Payroll Accuracy. (2023). *IJCSPUB - International Journal of Current Science* (www.IJCSPUB.org), ISSN:2250-1770, 13(2), 237–256, May 2023. Available: https://rjpn.org/IJCSPUB/papers/IJCSP23B1502.pdf
- Saha, Biswanath, and A. Renuka. (2020). Investigating Cross-Functional Collaboration and Knowledge Sharing in Cloud-Native Program Management Systems. *International Journal for Research in Management and Pharmacy, 9*(12), 8. Retrieved from www.ijrmp.org.
- Edge Computing Integration for Real-Time Analytics and Decision Support in SAP Service Management. (2025). *International Journal for Research Publication and Seminar, 16*(2), 231–248. https://doi.org/10.36676/jrps.v16.i2.283