

Jayas Piya

✉ jayaspiya52@gmail.com ☎ +977-9865005541 📍 Kathmandu, Nepal 🌐 /jayaspiya

🌐 /jayaspiya 🔗 <https://jayaspiya.com.np/>

Profile

Data Professional with 4+ years of experience designing and delivering large-scale **healthcare data platforms, analytics solutions, and enterprise data integration systems** using cloud and modern data engineering technologies. My research interests lie in trustworthy and scalable AI systems for real-world business applications, particularly Enterprise Agentic Systems, Retrieval-Augmented Reasoning, and Reliable Autonomous Workflows.

Work Experience

Abacus Insights Nepal

Jun 2025 – Present

Software Engineer II (R&D)

- Conducted early-stage R&D and adoption of generative AI, intelligent automation, and knowledge-assisted workflows, evaluating emerging approaches and their potential applications across enterprise business and data engineering processes.
- Evaluated internal LLM-based tools and agentic workflows prior to production release, assessing workflow effectiveness, reliability, and operational suitability for enterprise environments.
- Supported the Client Data Engineering Teams with enterprise data integration and interoperability, scalable data pipelines, automated documentation.
- Partnered with product and delivery teams to validate new capabilities and gather feedback, contributing to the development of reliable, user-centered solutions for enterprise-scale operations

Techkraft Inc

Software Engineer II

Jan 2024 – May 2025

- Engineered HL7 FHIR data integration framework, decoupling business logic from external data sources to ensure seamless enterprise-level interoperability.
- Automated data pipeline documentation, DAG flow visualization, and data lineage mapping, eliminating manual overhead and ensuring complete system visibility.
- Architected fail-safe data pipelines utilizing an Audit Balance Report framework and real-time observability metrics to guarantee data integrity and correct ingestion.
- Mentored junior engineers and led technical training sessions, contributing to the team's knowledge base to prevent technical silos and accelerate new hire onboarding.

Software Engineer

Oct 2022 – Dec 2023

- Orchestrated batch processing workflows, conducting performance profiling and query optimization to ensure efficient retrieval and lower compute costs.
- Automated the analysis process to solve the integration challenges of mapping raw client data (Bronze) to standard healthcare formats (Silver).
- Built stored procedures, UDFs, and modular queries to apply reusable healthcare logic, transforming raw staging data into optimized, analytics-ready tables.
- Improved overall team productivity and reduced manual effort by proactively automating repetitive technical tasks and debugging complex data pipeline issues.

Data Trainee

Jul 2022 – Sep 2022

- Completed the Data Engineering Fundamentals program, specializing in database design, ETL architecture, and modern data warehousing principles.
- Optimized internal operations by building Google Apps Script and Python automation tools, reducing manual data entry and streamlining cross-departmental reporting.

Education

BTEC HND, Computing

2019 – 2021

ISMt College

Grading: Distinction

BSc(Hons) Computing

2021 – 2022

Softwarica College of IT and E-commerce

Research Capstone: Retail Intelligence and Decision Support System

Overall Percentage: 80.25%

Projects

Healthcare FHIR Data Integration Framework [🔗](#)

Designed and developed a scalable healthcare data integration framework using PySpark to ingest, transform, and standardize HL7 FHIR clinical resources. Implemented Bronze-Silver-Gold data architectures and healthcare data models to create reliable, analytics-ready datasets from heterogeneous clinical records. Established reusable data processing pipelines to support downstream analytics, reporting, and machine learning applications.

Information Retrieval-Based Schema Mapper

Developed an information retrieval system to automate candidate schema mapping between source and target data models. Represented data dictionary metadata as searchable textual documents and applied TF-IDF vectorization with similarity-based ranking to identify semantically related fields. Designed the solution to reduce manual effort in early-stage ETL pipeline design, accelerating data integration workflows, and supporting enterprise data alignment.

30-Day Patient Readmission Risk Prediction Using Administrative Healthcare Data [🔗](#)

Developed a machine learning solution to predict 30-day hospital readmissions using longitudinal healthcare records and administrative patient data. Engineered predictive features from encounter histories, ICD-coded diagnoses, discharge outcomes, length of stay, chronic disease indicators, and healthcare utilization patterns. Built and evaluated XGBoost, Random Forest, and Logistic Regression models using cross-validation and hyperparameter optimization to compare predictive performance and support risk stratification for proactive patient care interventions.

Retail Intelligence and Decision Support System [🔗](#)

Developed a retail intelligence and decision-support platform that transformed operational data into actionable business insights through data warehousing and interactive analytics. Designed analytical data models and reporting pipelines to improve data accessibility, support performance monitoring, and enable evidence-based business decisions. Implemented secure and scalable backend database infrastructure to support concurrent analytical workloads and enterprise reporting requirements.

Certificates

- dbt Fundamentals - dbt Labs
- Google Data Analytics - Coursera
- Big Data Fundamentals with PySpark - Datacamp
- Oracle Database 12c: Basic SQL - LinkedIn
- Oracle Database 12c: Advance SQL - LinkedIn
- Software Engineering Principles Python - Datacamp