

Priyanshu Semwal

📍 Gurugram, India | 🌐 github.com/Priyanshu2425 | [in linkedin.com/in/semwalpriyanshu](https://www.linkedin.com/in/semwalpriyanshu)

SUMMARY

Data Scientist and AI Engineer with 1+ year of production experience in enterprise financial systems (JP Morgan) and automated AI pipelines. Proven track record of developing applied ML models and LLM integrations that drive sustained capital and revenue growth, including an intelligence platform generating ₹1.5L/month in revenue. Strong backend foundation in Python, advanced SQL, AWS serverless pipelines, and AI architecture; actively upskilling in ML via Scaler.

EXPERIENCE

MARXX.AI

Mar 2026 – Jun 2026

AI Data Scientist / Data Engineer

- Clients: Spinny, Kapiva, Hero, Nerolac, Banksy Museum (USA) + 15 others
- Architected a full-stack AI-powered ad intelligence SaaS platform; FastAPI + React 18/TypeScript frontend serving 20+ clients including Spinny and Hero, generating ₹1.5L/month in recurring revenue
- Integrated 5 LLM providers (OpenAI GPT-4, Anthropic Claude, Google Gemini) and computer vision models; shipped semantic search, AI creative analysis, and conversational intelligence in production
- Designed serverless AWS data pipelines (Lambda, SQS, Step Functions) to ingest and process unstructured data from 500+ brands daily, implementing robust schema definitions and system logging for high reliability
- Engineered a multi-modal ML enrichment pipeline with computer vision, vector embeddings, and GPU-enabled Docker workers; surfaced audio, visual, and text signals from ad creatives

Tata Consultancy Services — Client: JP Morgan

Apr 2025 – Mar 2026

Data Engineer / Systems Engineer

- Engineered and maintained 2 production Java payment flows within TCS BaNCS Corporate Actions, supporting JP Morgan's custody platform settling \$1T in securities daily across 100+ markets
- Engineered XML/JSON data pipelines and complex SQL queries for data migration across 3+ releases. Maintained strict data integrity under financial-grade SLA constraints for a custody platform settling \$1T in securities daily
- Diagnosed and resolved a high-severity data mapping bug; defect would have broken JPM's downstream integration in UAT — root cause identified, fix approved by stakeholders and deployed on deadline
- Primary developer on assigned flows within an Agile team of 4; releases certified by JP Morgan's product team targeting >90% STP across 4M+ corporate action events monthly

Recur Club

Oct 2024 – Apr 2025

Machine Learning Intern

- Built a document-processing pipeline for AICA, Recur Club's AI Credit Analyst engine, to handle Aadhaar, PAN, and other face-bearing financial documents submitted during borrower underwriting
- Built a document classification system using embeddings and a supervised learning model, achieving 99% precision classifying financial documents across 22 document types for the credit underwriting pipeline
- Trained a Random Forest classifier on historical data to flag borrowers likely to miss upcoming EMI payments, directly surfacing early repayment-risk signals to optimize the credit underwriting workflow
- Worked across a Java backend and Pandora, the team's Python backend, integrating ML models into production via message queues

PROJECTS

- **GenAI Market & Risk Intelligence Pipeline:** Architected an automated data pipeline to parse unstructured web data in real-time using Jina Reader. Deployed a memory-optimized Gemma 3 LLM architecture to extract macroeconomic sentiment and classify emerging financial risks, enabling dynamic, data-driven strategy adjustments.
- **Predictive Credit Default Engine:** Developed a robust machine learning model to assess credit default probability. Engineered complex, optimized SQL queries for time-series feature extraction from massive datasets. Implemented

comprehensive Python logging and modular architecture to ensure the prediction pipeline met production-grade enterprise standards.

- **Customer Segmentation & LTV Forecaster:** Processed large-scale transaction datasets using statistical clustering techniques to segment user behavior. Forecasted Customer Lifetime Value (CLTV) to optimize marketing spend allocation, aligning technical data outputs with long-term, endogenous capital growth strategies.

SKILLS

Languages & Frameworks: Python, SQL, Java, JavaScript, TypeScript · FastAPI, Spring Boot, React 18, Next.js, Node.js, REST APIs, JWT, Microservices

AI / ML: Applied ML, Random Forest, Statistical Clustering, Document Classification, LLMs, OpenAI GPT-4, Google Gemini, Anthropic Claude, Gemma 3, Computer Vision, Vector Embeddings, Semantic Search, Multi-modal ML Pipelines

Cloud & Databases: AWS (Lambda, SQS, Step Functions, EventBridge, CDK), Docker, Serverless · PostgreSQL, MongoDB, Redis

Tools & Practices: Git, Agile, Data Pipelines, Message Queues, System Design, STP · DSA in Java

CERTIFICATIONS

AWS Certified AI Practitioner (AIF-C01) — Amazon Web Services

AWS Certified Cloud Practitioner (CLF-C02) — Amazon Web Services

EDUCATION

Guru Gobind Singh Indraprastha University

2020 – 2024

B.Tech, Computer Science & Engineering GPA: 8.9 / 10