

SAMBIT BEHURA

AI / GenAI Engineer · Data Scientist

Bhubaneswar, Odisha, India | sambitkx2020@email.iimcal.ac.in | linkedin.com/in/behura-sambit | github.com/sbehu

SUMMARY

AI/GenAI engineer and data scientist with 9+ years of professional experience — including 6.5+ years across McKinsey, American Express, and consumer lending — now focused on productionizing agentic AI systems. Built and deployed multi-agent RAG pipelines (CrewAI, LangChain, Pinecone/ChromaDB) on AWS ECS Fargate with CI/CD and LLM-as-judge evaluation frameworks, on top of a core background in ML-driven decisioning that has driven ~\$200M in incremental business impact through XGBoost/logistic-regression models, causal experimentation, and credit risk analytics.

EDUCATION

Post Graduate Diploma in Business Analytics (PGDBA) — IIM Calcutta, ISI Kolkata & IIT Kharagpur (2018 – 2020)

B.Tech, Electrical and Electronics Engineering — National Institute of Technology, Tiruchirappalli (2011 – 2015)

SKILLS

GenAI / LLM: Multi-agent orchestration (CrewAI), LangChain, RAG architecture, Pinecone, ChromaDB, hybrid retrieval (dense + BM25 + cross-encoder rerank), prompt engineering, OpenAI Structured Outputs, LLM-as-judge evaluation (RAG Triad)

ML / Data Science: Python, XGBoost, Random Forest, Logistic Regression, SHAP, feature engineering, hypothesis testing, A/B / test-control experimentation, GINI/PSI/CSI model monitoring, credit risk & scorecard modeling

MLOps / Cloud: AWS (ECS Fargate, ECR, S3, CloudWatch), Docker, GitHub Actions CI/CD, Streamlit, SQL

Certifications: Building and Evaluating Advanced RAG · LangChain for LLM Application Development · Machine Learning A-Z · Credit Risk Modelling in Python

INDEPENDENT GENAI PROJECTS

Chargeback Intelligence Agent — Compliance-grade RAG pipeline for card-network dispute resolution

github.com/sbehu/chargeback-intel-agent

- Built a deterministic, auditable RAG pipeline grounding chargeback dispute defenses in actual Visa/Mastercard rule text via query expansion, hybrid retrieval (Pinecone dense + BM25 + FlashRank cross-encoder rerank), and OpenAI Structured Outputs for schema-constrained, Pydantic-validated verdicts.
- Built an LLM-as-judge RAG Triad evaluation harness against 400 real, stratified CFPB consumer-complaint cases, reaching 0.97 answer relevance / 0.75 groundedness / 0.74 context relevance on the labeled sweep, with full audit-trail logging to S3.
- Engineered header-aware PDF rulebook chunking (via pdfplumber) that splits at detected rule-condition boundaries instead of fixed-word windows, eliminating cross-rule contamination; deployed on AWS ECS Fargate via Docker/uv with a GitHub Actions CI/CD pipeline.

AI Financial Analyst Platform — Multi-agent RAG system for 10-K financial analysis

github.com/sbehu/financial_agent_intel

- Orchestrated a multi-agent workflow (CrewAI) to parse and analyze 6 years of American Express 10-K filings, dynamically benchmarking performance against live Visa/Mastercard competitor data pulled from external APIs.
- Solved fiscal-year metadata blending and financial-table hallucination by isolating vector chunks by filing year and pairing structural table chunking with a dedicated math agent for numeric verification.
- Built an automated RAG Triad evaluation layer against a 200+ question synthetic golden dataset with a Streamlit dashboard, and deployed via Docker on AWS ECS Fargate through GitHub Actions CI/CD.

EXPERIENCE

Senior Consultant — EXL

Mar 2026 – Apr 2026

- Built fraud detection models for a Long Term Care (LTC) insurance client, leveraging claims behavioral data to flag anomalous claim patterns.

Assistant Manager, Global Credit Services — American Express

Jun 2022 – Nov 2025

- Progressed from the Client Onboarding & Digital Capabilities team into the AI & Modeling (AI&M) team over a 3.5-year tenure, moving from operational analytics into applied ML and experimentation.
- Built an XGBoost-based B2B product recommendation model (GINI 0.85) to identify optimal buyer-supplier matches from accounts-payable data, engineering 70+ features with cross-validation and hyperparameter tuning, and improved stakeholder trust via SHAP-based explainability.
- Quantified ~\$200M in annual incremental billed business by measuring marketing campaign impact through randomized test-control experimentation and hypothesis testing, directly informing leadership's campaign investment decisions.

Senior Manager, Analytics — SMFG India Credit (Fullerton India)

Dec 2021 – Jun 2022

- Monitored and validated Application, Behavioural, and Collection credit risk models using GINI, PSI, and CSI stability metrics, identifying performance drift and recommending recalibration.

Data Scientist — U GRO Capital

Apr 2020 – Nov 2021

- Built U GRO's first-generation business-loan application scorecard, engineering 1,000+ features from bank statement and bureau data and shipping a logistic regression model (GINI ~0.80) that became the core of credit underwriting.

Data Science Intern (Pre-Placement Offer) — McKinsey & Company

Oct 2019 – Mar 2020

- Developed a personal-loan cross-sell model targeting early-tenure consumer-durable and two-wheeler loan customers using XGBoost and Random Forest, improving targeting effectiveness.

EARLY CAREER (~3 YEARS)

Assistant Professor of Mathematics, FIITJEE Limited, Bhubaneswar (2016 – 2018) · Business Analyst (UI Developer), Royal Bank of Scotland, Chennai (2015)