# DEBANJAN DEY

ddaccount26@gmail.com | LinkedIn | Website | +91-8893541023

## SUMMARY

Senior Data Analyst with 5+ years' expertise in data analysis, risk management, and automation across healthcare and finance. Experienced in CI/CD pipelines and ETL processes using tools like SQL and Python, and visualization with Qlik and Tableau. Former Financial Crime Risk Analyst at HSBC, building data-driven solutions to mitigate financial crime. Currently at Phreesia Healthcare, optimizing patient checkins through analytics and automation. Skilled in machine learning for compliance, pattern detection, and fraud prevention. AI enthusiast with hands-on experience using LLM APIs, prompt engineering, AI tools, agents, and Retrieval-Augmented Generation (RAG) systems.

## WORK EXPERIENCE

**Senior Data Analyst, Phreesia**                                                            **Jun 2024 – Present**
- Built a UI-driven automation tool that generates customized Business Review decks for 4000+ clients, auto-fetching data from Qlik and Excel to populate PowerPoint slides — saving **3 FTEs** through full-cycle reporting automation.
- Developed a **Streamlit**-based chatbot using **Groq API**, enabling internal stakeholders to query and interact with Confluence documentation in real time, reducing follow-up queries and saving an estimated **20+ hours per month** in manual support.
- Designed and implemented a HIPAA-compliant data tokenization architecture for third-party measurement workflows, integrating Snowflake, Streamlit and vendor specific tokenization pipelines to securely pre-process, encrypt, and deliver PHI data while maintaining end-to-end availability and encryption controls.

**Senior Data Analyst, HSBC**                                                                **Nov 2022 – Jun 2024**
- Developed an in-house UI tool for SWIFT MT/ ISO 20022 test file generation, automated alert reconciliation for model monitoring of the transaction screening tool. It replaced a third-party tool, yielding an annual cost saving of ~**$1 million**.
- Implemented a post detection layer for payments screening tool, resulting in up to a **7% reduction** in false positive alert volume. It uses text matching on fuzzy logic Levenshtein distance, risk score assignment through **K-means clustering** for identifying low-risk patterns to create whitelists for payment filter.
- Developed a machine learning model as an add-on to the post-detection layer for name screening engine that reduced the false positive alert volume by **20-35%** before feeding into the alert discounting engine.

**Data Analyst, HSBC**                                                                        **Aug 2020 – Oct 2022**
- Built a Python tool to validate regulatory list updates across regions, automating reconciliation and reducing manual effort by 80%, saving **2 FTEs** in model monitoring.
- Catered to ad-hoc queries for alert data extraction from Customer and Payments **Data Lake** using **Hive SQL**, development and maintenance of alert dashboards in Tableau and Qlik Sense.
- Automated periodic validation of the name screening engine using exact and fuzzy matches of sanctioned entities, AML, and PEPs from World-Check, plus alert reconciliation, reducing risk and ensuring regulatory-compliant model effectiveness.

## PROJECTS

**AI Agent Evaluation & Benchmarking Specialist | Turing (Contract for Amazon AI)**
- Engineered complex, programmatically verifiable agentic task scenarios for Amazon's AI agent evaluation via tau-bench, architecting multi-step, multi-API workflows across enterprise domains (finance, payroll, HR, incident management) where agents execute 15-23+ sequential tool interactions to query databases, retrieve state via API calls, and generate deterministic outputs.

**Career Nexus Academy** | https://careernexusacademy.in/
- Spearheaded the complete development and deployment of Career Nexus Academy from inception to production, architecting end-to-end full-stack infrastructure including front-end interfaces, back-end APIs, database design, and cloud deployment while executing all phases from system design and feature implementation through quality assurance, DevOps configuration, and production release.

## EDUCATION

| Degree/Program | Institution | Graduation Year | CGPA/Percentage |
| --- | --- | --- | --- |
| M. Tech in Power Systems | IIT Delhi, New Delhi | 2020 | 8.67/10 |
| B. Tech in EEE | NIT Calicut, Kerala | 2017 | 8.41/10 |
| Class XII | Ramkrishna Mission Vidyalaya, Viveknagar, Tripura | 2013 | 93.6% |
| Class X | Ramkrishna Mission Vidyalaya, Viveknagar, Tripura | 2011 | 10/10 |

## TECHNICAL SKILLS

**AI/ML**: Logistic Regression, XGBoost, Tensorflow, scikit-learn, LLM APIs (OpenAI, Groq, Huggingface), Langchain, AI Agents, RAG
**Data Tools**: Python, SQL, Apache Hive, Big Data, Snowflake, Streamlit, BigQuery, ETL Pipelines, MongoDB, Tableau, Qlik Sense
**Dev & Collaboration Tools**: Flask, HTML, CSS, Javascript, Git, Vercel, JIRA, Confluence

## CERTIFICATIONS

Lean Six Sigma Yellow Belt - Catalyst Consulting [Credential ID: 88613829]

Supervised Machine Learning: Regression and Classification - Stanford Online [Credential ID: FTWY79DQL5AC]

Advanced Learning Algorithms - Stanford Online [Credential ID: 18V1G7BVB6OI]

Intro to Deep Learning & Neural Networks with Keras - Coursera [Credential ID: XC4W83JJEVXS]