

# HRIDAY SAHA

## SENIOR DATA SCIENTIST | FRAUD RISK & NLP SYSTEMS | APPLIED MACHINE LEARNING

Bangalore, India | +91 8670745666 | job@hridayasaha.com | linkedin.com/in/ihridayasaha | hridayasaha.com

### PROFESSIONAL SUMMARY

Senior Data Scientist with 4+ years in fraud detection, NLP automation, and risk analytics across global enterprises. Delivered \$1.7M+ in business impact using Sentence-BERT, XGBoost, and advanced feature engineering, collaborating with engineering teams to deploy production-ready ML solutions.

### CORE COMPETENCIES

- **Machine Learning & AI:** Supervised & Unsupervised Learning, Ensemble Methods (XGBoost, LightGBM), Fraud Detection & Anomaly Detection, NLP (Transformers, Sentence-BERT, SpaCy), Neural Networks & Deep Learning
- **Programming & ML Stack:** Python, SQL, R, Scikit-learn, TensorFlow, PyTorch, Hugging Face, SpaCy
- **Data Science & Analytics:** Predictive Modeling & Feature Engineering, A/B Testing & Hypothesis Testing, Time Series Forecasting
- **Visualization & BI:** Power BI, Tableau, Plotly, Matplotlib, Seaborn, Data Storytelling, Dashboard Design
- **Infrastructure & Workflow:** Python-based ETL Pipelines, Model Evaluation & Validation Frameworks, Monitoring Metrics & Data/Concept Drift Analysis, Collaboration with Engineering & DevOps, Git, Docker (working knowledge)
- **Collaboration & Strategy:** Cross-functional Stakeholder Management, Technical Mentoring, Translating ML Outcomes into Business Impact

### PROFESSIONAL EXPERIENCE

*Dec 2024 – Present*

#### **Data Science Trainer & Consultant | The Sparks Foundation, Bengaluru, India**

- Designed and delivered **45+ hours of applied data science training** to 150+ mid-career professionals via SGUnited partners (Rise by BCG, Lithan Academy, NTUC Learning Hub, Kaplan Inc.), achieving **92% completion** and **85% job-readiness improvement**.
- Developed industry-aligned ML curriculum covering **Python, Scikit-learn, XGBoost, Transformers, and Power BI**, with hands-on projects in fraud detection, NLP pipelines, and predictive modeling.
- Mentored **60+ capstone projects** across customer segmentation, time-series forecasting, and sentiment analysis, with **40+ reaching production-ready design standards**.
- Provided technical consulting for **3 SME clients**, implementing Python-based ETL workflows and executive Power BI dashboards, reducing manual reporting by **40%** and improving decision-making speed by **30%**.

*Oct 2023 – Sep 2024*

#### **Data Scientist | AXA Group Operations, Paris, France**

- Built an **NLP-driven semantic similarity system** combining BM25 retrieval, fine-tuned Sentence-BERT ranking, and LightGBM scoring to automatically map local audit issues to SAF controls.
- Achieved **93% Top-1** and **99% Top-3 mapping accuracy**, reducing manual audit effort by **96%** and generating approximately **\$180K in annual savings**.
- Collaborated with auditors, security assurance specialists, data architects, and ML engineers to translate domain requirements into scalable ML solution designs.
- Designed **explainable model outputs** (confidence scores, key-phrase highlights) to improve auditor trust and adoption across multiple stakeholder groups.
- Worked closely with engineering and DevOps teams to **define deployment requirements**, retraining strategies, and data drift monitoring for enterprise ML workflows.

Oct 2022 – Sep 2023

### Data Governance & Analytics Project Manager | AXA Group Operations, Paris, France

- Designed and validated a **real-time fraud detection solution** using XGBoost ensembles with engineered features including transaction velocity, geo-mismatch, device fingerprints, and graph-based merchant–user linkages.
- Achieved **94.2% F1-score**, supporting a platform handling **1M+ daily transactions** with **sub-100ms inference requirements**.
- Applied **time-series cross-validation** techniques to address fraud pattern drift and ensure model robustness over time.
- Led a team of **3–5 data professionals**, coordinating with **40+ stakeholders** across fraud, risk, compliance, and data governance functions.
- Delivered a **\$1.2M analytics portfolio** with **98% on-time delivery**, contributing to **\$1.5M+ in annual fraud loss prevention**.

Sep 2021 – Sep 2022

### Data Analytics Associate | AXA Group Operations, Kuala Lumpur, Malaysia

- Developed classification models and statistical validation pipelines (t-tests, ANOVA) to support enterprise risk analysis workflows.
- Automated Python-based ETL and anomaly detection processes, reducing data discrepancies by **80%**.
- Built interactive **Power BI dashboards** with live KPIs, improving business decision effectiveness by **10%**.
- Managed a team of **3 analysts**, mentored **2 junior analysts**, and collaborated with multiple cross-functional teams to streamline analytics delivery.

## EDUCATION & CERTIFICATIONS

---

2021

**Master of Science in Data Science and Business Analytics**  
*Asia Pacific University of Technology and Innovation (APU), Malaysia*

2018

**Bachelor of Technology in Computer Science and Engineering**  
*MCKV Institute of Engineering, India*

### Professional Certifications:

- Microsoft Power BI Analytics (2022)
- SAS Data Science Professional (2021)
- IBM Data Science Professional Certificate (2020)

## AWARDS & RECOGNITION

---

- **AXA Innovation Excellence Award (2023)**: Recognized for building an ML-based audit automation system delivering \$180K+ annual savings.
- **AWS Build Malaysia Semifinalist**: Top 10% among 500+ participants for a cloud-based ML data processing solution.
- **Best Rotaract President (2017–18)**: Led 80+ events, co-organized 25+, and delivered guest talks across districts.

## ADDITIONAL INFORMATION

---

- Fluent in English, Bengali, and Hindi
- Open to relocation and remote opportunities