

HARSH VARDHAN SINGH

Senior Data Scientist

@ harshsinghv@gmail.com

📞 9414583715

📍 Jaipur, Rajasthan - India

in linkedin.com/in/harshvardhan33

WORK EXPERIENCE



Senior Data Science Consultant

📅 July 2024 - Present

📍 Jaipur, India

- Established a **Microsoft Azure and Databricks-focused Data Science practice** under the Intelligent Automation service line, scaling from **1 to 15+ members** in a year.
- Successfully closed multiple Data & AI presales engagements across APAC and Europe, contributing to a **multi-million-dollar** revenue pipeline, while also creating and driving **GTM motions** to accelerate solution adoption and customer acquisition
- Successfully delivered **5+ end-to-end Data Science projects** in BFSI, Energy, and Healthcare, leveraging Azure AI, Databricks, and MLOps to drive measurable business impact.
- Spearheaded the development of **Azure-based AI/ML accelerators**, strengthening Xebia's Azure and Databricks offerings.
- Led strategic initiatives that enabled Xebia to secure **Microsoft's Global Data & AI Partner Specialization**, enhancing credibility and market positioning within the Microsoft ecosystem.



Senior Data Scientist

📅 July 2022 - July 2024

📍 Jaipur, India

- Developed and designed industry-leading AI solutions in collaboration with **Microsoft Research in Energy, Sustainability and Oil & Gas Domain**.
- Expertise in leveraging AI and BigData solutions to enhance, automate, and optimize business processes.
- Developed **cloud native frameworks** and solution accelerators for rapid AI integration.
- Achievements** : As a technical lead, I have overseen the delivery of **10+ projects** and actively contributed to acquiring new clients through **technical pre-sales** efforts. Furthermore, I played a crucial role in creating our company's first product **and the SaaS offering**.
- Led a team of 10 data scientists to drive impactful data-driven initiatives.



Data Scientist

📅 May 2021 - July 2022

📍 Gurugram, India

- Worked on **building and designing** end to end **AI based solutions** to improve overall processes in logistics and supply chain.
- Developed solution for **QC-Reverse** Pickup shipment which saves upto **3Cr annually** and improve QC process.
- Historical data distribution based Customer TAT Re-Promising Engine on OD pair level for DC and Hubs.
- Guided and supported multiple trainees across various projects to develop data science skills and deliver value to business.

DAIICT

Research/Teaching Assistant - Machine Learning and Deep Learning

📅 June 2019 - May 2021

📍 Gandhinagar, India

- Assisted **300+ students** in coding for lab sessions.
- Mentorship of **20+ students** in Machine Learning and Deep Learning. Supervised and conducted multiple lab practical in each semester.

PROJECTS & STRATEGIC INITIATIVES

AI Enabled Regulatory Petition Drafting System

- Designed and developed a **RAFT (RAG + Fine-tuning)** solution to automate and streamline **regulatory petition drafting** by extracting and synthesizing unstructured legal data, addressing inefficiencies, errors, and delays in manual processes.
- Tech Stack** : Azure Document Intelligence, Azure AI Search, Azure AI Foundry, Azure OpenAI Models.
- Business Impact** : Achieved up to **60% reduction** in drafting time by minimizing manual effort involved in the iterative generation of petition drafts.

Generative AI-Driven Recipe Ingestion and Validation for ARAS PLM

- Designed and implemented a GenAI based automation solution integrated with ARAS PLM for intelligent data extraction and validation to streamline product lifecycle processes.
- Combined RAG techniques with extensive prompt engineering and domain-specific business rules to ensure accurate and validated information flows into the PLM system.
- Tech Stack**: Azure AI Foundry, Azure Document Intelligence, Power Platform, Power Automate.
- Business Impact**: Enhanced document validation efficiency and accuracy in PLM workflows, reducing manual effort and turnaround time for engineering change orders and documentation approvals.

Unified AI Platform for Financial Intelligence for a Holdings Company

- Led the design and development of an **integrated AI platform to unify siloed AI use cases across the organization**, enabling enterprise-wide AI democratization and adoption.
- Catered Multiple usecases - **IC Memo, Corporate Memorandum Structuring, AI Procurement Agent** using Azure AI.
- Integrated **Microsoft Fabric** to support complex financial modeling pipelines and provide real-time analytics and decision support across business units.
- Tech Stack**: Azure AI Foundry, Microsoft Fabric, Azure OpenAI, Azure Synapse, Custom UI with Open Source Frameworks.
- Business Impact**: Enabled cross-departmental access to AI capabilities, reduced model redundancy, and accelerated time-to-insight by consolidating fragmented AI efforts into a single, scalable platform.

Generative AI Presales Lifecycle Optimizer - Internal Solution

- Designed and implemented an internal GenAI solution to automate the presales lifecycle, covering RFP generation, slide creation, proposal drafting, and architecture recommendations using historical organizational knowledge.
- Modeled a vector database to **semantically index** past proposals, client decks, operating models, and architectural patterns, enabling **intelligent retrieval** via a **RAG-based workflow** powered by Azure OpenAI.
- Tech Stack** : Azure OpenAI, Azure AI Search, Azure Cosmos DB, LangChain, Python, Aspose, python-pptx.
- Business Impact** : Achieved up to **40% reduction** in RFP response time cutting down the typical 1-2 week cycle to 5-7 days in the V1 rollout.

Power and Utility Forecasting Framework

- Collaborated with **Microsoft Research** to create a cutting-edge energy sector solution accelerator for generation and demand forecasting through the incorporation of micro-climatic external factors.
- Tech Stack** : Azure Machine Learning, Deep Learning, Pytorch
- Business Impact** : Accurate Forecasting, Penalty Minimization for PPA clients, Effective Maintenance lifecycle, Reduction in overall solution lifecycle and upto **40% reduction** in penalties.

Vision Analytics : Energy, Oil and Gas

- Developed a cutting-edge computer vision framework tailored to the **Energy Sector**, focusing on the transmission and distribution components of the value chain.
- Implemented various **point solutions** built upon the framework, including EL Panel Defect Detection, Tower Insulator Damage Detection and Segmentation, Corrosion Detection, EHS Copilot etc.
- Tech Stack** : Azure Machine Learning, Deep Learning, PyTorch, Vision Transformers, YOLO V8, Azure Vision Studio and Florence.
- Business Impact** : Automated event driven inspection, failure prevention, network optimization.

EDUCATION

M.Tech (ICT) in Machine Learning

Dhirubhai Ambani Institute of Information and Communication Technology - 7.59 CGPA

📅 July 2019 – July 2021 📍 Gandhinagar, Gujarat

B.Tech (Computer Science Engineering)

JECRC University - 9.02 CGPA

📅 2015-2019 📍 Jaipur, Rajasthan

TECHNICAL SKILLS

- Machine Learning, Deep Learning, Statistics, Generative AI and LLMs
- Language - Python, SQL
- Cloud - AWS, Azure
- Database Technologies - MYSQL, MongoDB
- Bigdata Frameworks - Spark, AWS EMR, Databricks.
- Docker, Kubernetes, Flask, Django, Redis, Tensorflow, Pytorch, OpenCV

MICROSOFT & DATABRICKS ACCOLADES



- Databricks Certified: Machine Learning Professional
- Databricks Certified: Generative AI Engineer Associate
- Databricks Certified Associate Developer for Apache Spark 3.0



- Microsoft Certified: Azure Data Scientist Associate
- Microsoft Certified: Azure AI Engineer Associate
- Microsoft Certified: Azure AI Fundamentals
- Microsoft Certified: Azure Data Fundamentals
- Microsoft Certified: Azure Fundamentals

RESEARCH WORK



Transform Domain Image Super Resolution using Deep Neural Network

- Research Work : Deep Learning architecture for **Image Super Resolution** using Transform Domain image frequency analysis. Paper published in **IEEE Journal - TNNLS 2021**
- [Read the paper here](#)

PROJECTS & STRATEGIC INITIATIVES

Enterprise Knowledge Advisor

- Implemented an Enterprise Knowledge Advisor that enables natural language powered **access and retrieval of information** making it easier for them to utilize information for better decision-making and operational efficiency.
- **Tech Stack** : OpenAI, Form Recognizer, and Azure Cognitive Search to process and index diverse documents, including text and images. Use of embedding models for document and query representations.
- **Business Impact** : Process Automation, Enhanced Decision Making, Customer Satisfaction etc.

AI Based QC-RVP

- Developed AI based solution for **QC-Reverse Pickups** in Logistics using AI. Solution ensures that a reverse pickup collected back from consignee is the actual product that needs to be picked up.
- **Tech Stack** : AWS Reko, Clustering, BERT, Django, MongoDB Integration, Asynchronous (Celery and Redis).
- **Business Impact** : Saved upto **3Cr** annually. Automated the process for QC reverse pickups by reducing the manual inspection load by almost 35%.

Enterprise Level Face Recognition Attendance System

- Co-Developed an enterprise-level face recognition attendance system for distribution centers (DCs)
- **Tech Stack** : Facenet Embeddings, Tensorflow, Django, MongoDB.
- **Business Impact** : Cost Effective, Prevents bad hires , Touchless Sign-in.

Customer TAT Repromise Engine in Logistics

- Developed a pipeline for **EDD Breached shipments** using **historical cumulative distributions** on OD pair level. **Multi-level mapping** to handle network changes on DC and Hubs.
- **Tech Stack** : AWS EMR, Pyspark and MYSQL.
- **Business Impact** : Customer Satisfaction, better SLA's, increased sales, better rescue plan and business compliance at Operations level.

Other Projects

- **24x7 Energy Matching Framework** - Renewable Energy Supply and Demand Mix Platform
- **Decision Management Framework** - Energy Arbitrage and MicroGrid Optimization,
- **Predictive Maintenance for Transformers** - Classifying Transformer Health Status
- **Distribution Network Control Tower** - Smart Meter Analytics
- **EV Load Identification** - find hotspot locations for high EV loads based on smart meter data.