

AHMAD MUJTABA

GenAI Engineer | Data Scientist

📞 +91-98070-52640 | 📩 ahmad.iiitk@gmail.com | 💬 LinkedIn | 🐾 GitHub

PROFESSIONAL SUMMARY

GenAI Engineer & Data Scientist at Deloitte US-India with 3+ years of experience delivering enterprise-grade AI solutions for US Healthcare and Automotive domains. Expert in architecting Reasoning-Driven IDP pipelines, elevating extraction accuracy from ~90% to ~99% by orchestrating Gemini-3-Pro workflows over GPT-5.1 baselines. Proficient in building Autonomous Computer Using Agents (OpenAI), boosting task success from ~38% to ~80% via Milvus-backed RAG orchestration for SOP grounding. Strong technical foundation in Azure Databricks ETL (Silver/Gold curation), classical ML (Scikit-learn/TensorFlow), and Dockerized deployment, translating complex unstructured data into measurable business value.

TECHNICAL SKILLS

Generative AI & LLMs: Agentic Workflows (Computer Using Agent), Model Context Protocol (MCP), Vision-Language Actions (VLA), RAG Orchestration (Milvus/Vector DBs), IDP (Intelligent Document Processing), Prompt Engineering (CoT, ReAct), Entity Extraction, Azure OpenAI (GPT-4/5), Google Gemini, LangChain, LangGraph, AWS Lex, LUIS, CLU

Machine Learning & Data Science: Predictive Modeling (Regression, Classification, Clustering), Computer Vision (CNN, PaddleOCR, AX Tree Parsing, Layout Analysis), NLP (NLTK, LSTM, RNN, Transformers), Ensemble Methods (XGBoost, Random Forest), Transfer Learning, Feature Engineering, EDA

Data Engineering: Azure Databricks (PySpark, Notebooks), ETL/ELT Pipelines, Data Warehousing (Medallion Architecture - Silver/Gold), Data Modeling, Complex Joins, MS SQL Server, Amazon S3

Cloud & DevOps: Azure (App Services, Cognitive Services, Key Vault, Blob Storage, ML Studio); AWS (Lambda, S3, CloudWatch, SageMaker); Docker (Containerization), Git/GitLab, CI/CD, Power BI

Libraries & Tools: Python, SQL, Scikit-learn, TensorFlow, Spark MLlib, Pandas, NumPy, Matplotlib, Seaborn

Domain & Soft Skills: Automotive, Manufacturing, FMCG, Warranty Analytics, B2B Reordering, Problem-Solving, Team Coordination, Cross-Functional Collaboration, Time Zone Management

EXPERIENCE

Deloitte US-India

GenAI Engineer | Data Scientist

Gurugram, India

July 2025 – Present

- Architected High-Precision IDP Pipeline: Engineered an end-to-end intelligent document processing system for unstructured insurance policies, integrating PaddleOCR for layout-aware extraction and LLMs for semantic reasoning over complex policy clauses.
- Achieved 99% Extraction Accuracy: Elevated entity extraction reliability from ~90% (GPT-5.1 baseline) to 99% by migrating to Gemini-3-Pro-Preview and implementing a "Canonical Comparison" self-correction loop to resolve ambiguity in validation logic.
- Architected Autonomous CUA Platform: Fast-tracked the deployment of OpenAI's Computer-Using Agent ahead of schedule by engineering a headless Linux Docker environment on AWS, integrating a custom frontend that streams the OS display for real-time "Human-in-the-Loop" observability.
- RAG-Driven Reliability Engineering: Boosted agent task success from 38% to ~80% by implementing a Milvus-backed RAG layer, retrieving granular Standard Operating Procedures (SOPs) to ground the agent's reasoning and reduce stochastic behavior.
- Vision Pipeline Optimization: Accelerated the perception-action loop and reduced token costs by refactoring the ingestion layer to leverage compressed JPEG artifacts, lowering latency without compromising UI element detection accuracy.
- Architected Playwright MCP Agent: Engineered a web automation tool using the Model Context Protocol (MCP) to interface LLMs with Playwright. Optimized context efficiency by parsing **Accessibility (AX) Tree snapshots** instead of raw DOM, reducing token usage by ~40% while improving element resolution accuracy.

- Developed Scalable ETL Frameworks: Built robust data ingestion scripts on Databricks using PySpark and Python; handled complex schema mapping, primary key reconciliation, and multi-table inner joins to produce clean, consolidated datasets for downstream analytics.

Cognizant Technology Solutions

Noida, India

Associate Data Scientist

Sep 2022 – May 2025

- Automated Warranty Analytics: Developed predictive models using Random Forest on Azure Databricks (PySpark) to classify warranty claims, improving decision recall and reducing manual review cycles for a major automotive client.
- Deployed Enterprise Chatbots: Designed a B2B conversational AI assistant using AWS Lex and Azure OpenAI, incorporating logic to parse unstructured orders (quantity, flavor, SKU) and automate the reordering process via AWS Lambda.
- Optimized Cloud Infrastructure: Enhanced Python-based applications on Azure App Services, resolving timeout bottlenecks and implementing backend logic via AWS Lambda and S3 for scalable data handling.
- Visualized Business Insights: Created interactive Power BI dashboards to track model performance and KPIs, directly supporting data-driven decision-making for global cross-functional teams.
- Collaborated Globally: Managed version control via GitLab and coordinated with global teams to streamline Azure-based deployments, achieving higher on-time delivery rates.

AiEnsured

Remote, India

Machine Learning Engineer Intern

Jul 2021 – Aug 2021

- Built Computer Vision Models: Contributed to the development of Object Detection systems using CNNs, assisting senior engineers in code optimization and error analysis.
- Enhanced Model Performance: Implemented regression and classification algorithms for diverse datasets, conducting feature engineering that improved baseline prediction accuracy.

EDUCATION

Indian Institute of Information Technology (IIIT)

Kurnool, India

Master of Technology (M.Tech) in Data Analytics and Decision Sciences

Oct 2020 – Jun 2022

Relevant Coursework: Machine Learning, Deep Learning, NLP, Computer Vision, Statistics, Probability & Discrete Math.

Maulana Azad National Urdu University

Hyderabad, India

Bachelor of Technology (B.Tech) in Computer Science Engineering

Aug 2015 – Jun 2019

Relevant Coursework: Data Structures & Algorithms, Engineering Mathematics, Web Development.

CERTIFICATIONS

Machine Learning Specialization | DeepLearning.AI via Coursera

Jan 2023