

Aachal Kala

+91 7694074213 — aachalkala2001@gmail.com — linkedin.com/in/aachal-kala-1050a31a6/ — github.com/Aachal622

Professional Summary

- Experienced Data Scientist with 3+ years designing end-to-end ML and LLM pipelines, delivering scalable solutions across Healthcare, Finance, and Non-Profit sectors.
- Expertise in Generative AI, Prompt Engineering, and RAG using LangChain, GPT-4, and Azure Open AI.
- Proven track record in building robust ETL pipelines, optimizing model accuracy, and automating data workflows achieving up to 90% faster processing.
- Strong in Python, SQL, and cloud platforms (Azure, AWS); adept at data wrangling, feature engineering, and production deployment of ML systems.
- Recognized for technical leadership, cross-functional collaboration, and translating business needs into high-impact AI solutions.

Skills

Generative AI: LangChain, Large Language Model (LLM), Vector DB, RAG, Prompt Engineering

Data Science: Machine Learning, Supervised Learning, Natural Language Processing (NLP)

Soft Skills: Leadership, Team Management, Analytical Skills

Cloud: AWS, Azure Platform

Languages: Python, SQL

Libraries: Pandas, Scikit-learn, NLP, OpenCV, Flask

Database: MySQL Server, PostgreSQL

Others: Stream-lit, Docker, GIT, HTML

Professional Experience

Nagarro, Jaipur (Feb 2023 - Present)

Data Scientist

- Developed production-grade AI systems using OpenAI and LangChain, transitioning legacy software stacks into intelligent, automated solutions.
- Designed and deployed RESTful Flask APIs supporting real-time LLM requests, improving scalability and reducing API latency by 40%.
- Implemented advanced AI capabilities by integrating Azure Open AI and Azure Cognitive Services, delivering intelligent features tailored to business needs.
- Built Retrieval Augmented Generation (RAG) pipelines by combining vector stores with GPT models, improving document comprehension accuracy by 30%.
- Received client appreciation for active engagement in the project, through understanding of business requirements, and valuable suggestions provided.
- Technology Stack: Python, Lang-Chain, Flask API, Azure Open AI, Azure Cognitive Services, AWS, LLM, Pandas, Scikit-learn, SQL, Stream-lit, Prompt Engineering.

Advanced Technology Consulting Service, Jaipur (Jul 2022 - Feb 2023)

Associate Data Scientist

- Developed a web app for pharmaceutical companies to analyze market potential, brand volume, and access scores using drug insurance data.
- Built and maintained data-driven applications, integrating ML workflows for enterprise clients in Pharma and Healthcare verticals.
- Integrated Machine Learning modules with the product application and actively participated in code reviews and sprint planning.
- Automated unstructured data ingestion from PDFs and Excel using Python-based parsers, improving ETL throughput and reducing errors by 40%.

Education

Poornima Institute of Engineering and Technology, Jaipur, India (2018 - 2022)

Bachelor of Technology - Computer Science and Engineering; CGPA: 8.6

Projects

Wildlife Conflict Resolution Platform (Sept 2024 - Present)

Client: World Wildlife Fund

- Engineered NLP pipelines with Azure OpenAI and LangChain to extract insights from unstructured text, improving relevance ranking and search accuracy.
- Automated data ingestion via Python-based scraping and database integration, reducing input errors by 95% and improving ingestion speed by 90%.
- Scalable Data Architecture: Designed and implemented a secure, reliable, and scalable data management system on AWS.
- Technical Leadership: Designed cross-functional teams, providing technical guidance and oversight to ensure successful project delivery
- Technology Stack: Python, Lang-chain, GPT Models, AWS Services, Selenium, Beautiful Soup, Open AI, Google Geo-coding API.

Proposal Genius (Apr 2024 – Aug 2024)

Client: Nagarro

- Built an AI-based requirements extraction engine using GPT-4 and Azure Cognitive Search to automate proposal parsing and project planning.
- Utilized GPT-4, Lang-Chain, and Azure Cognitive Search to accurately extract project data such as dates, requirements, and challenges,
- Responsibilities: Managed AI tool integration, validated project requirements, and ensured performance standards.
- Technology Stack: Python, Flask, Lang-Chain, Open AI, GPT-4, LLM, Prompting, Gemini Model, Pandas, Azure Cognitive Search, Azure Blob Storage.

Voice Assistant App (Jan 2024 - Mar 2024)

Client: Ma-Panda

- Developed a voice-based educational tutor for children (ages 3-4) named “Mickey” using Flask and Azure Open AI’s GPT-4 model.
- Integrated Azure Cognitive Services for real-time speech processing, enhancing user interaction through voice-enabled AI tutoring.
- Implemented personalized, child-friendly responses and interactive quizzes for engaging educational experiences.
- Ensured secure and efficient upload and download operations with Azure Data Lake and Blob Storage.
- Responsibilities: Developed speech-to-text features, integrated AI models, and ensured content was age-appropriate and engaging.
- Technology Stack: Python, Flask, Azure Cognitive Services (Speech Service, Azure Open AI).

AI Knowledge Base Module (Jul 2023 – Dec 2023)

Client: The Family Office

- Developed a knowledge base module that facilitates document processing and question-answering functionalities using Azure services and Open AI’s GPT-35-turbo.
- Implemented document upload, processing, searching, and deletion in Azure Data Lake Storage. Integrated Open AI models for answering questions and generating embeddings.
- Ensured secure and efficient upload and download operations with Azure Data Lake and Blob Storage.
- Responsibilities: Implemented API endpoints, integrated Azure services, set up Open AI models, and ensured secure operations.
- Technology Stack: Python, Flask, Lang-Chain, Azure Services (Data Lake Storage, Key Vault, Blob Storage, Cognitive Services), Open AI (GPT-35-Turbo, text embedding-Ada), Pandas.

Healthcare Data Analytics Platform (Jan 2022 - Jul 2023)

Client: Indegene

- Healthcare Analytics Platform: Developed a scalable healthcare analytics platform for USA-based patient data, utilizing Python and Machine Learning algorithms.
- Created an archival and restoration pipeline for the uploaded data in S3 resulting in 70% AWS Cost reduction monitored over a span of 4 months.
- AWS Deployment: Successfully deployed the application on AWS, leveraging various services to ensure seamless functionality and scalability.
- Technology Stack: Python, AWS services (Lambda, DynamoDB, Simple Storage Service, Sage-Maker, Event-Bridge, Queue Service), Machine Learning.

Certifications

- Microsoft Azure AI Fundamentals – [AI-900]
- Data bricks: Generative AI Fundamentals – [Databricks Academy]
- Linked-In: Azure Essential Training for Developers
- Microsoft 365 Copilot First Look
- Linked-In: Introduction to Artificial Intelligence
- Coursera: Python Specialization