

SUMMARY

Data Scientist with 4 years of experience building AI/ML solutions, ETL Pipelines and Generative AI systems. Skilled in LLMs Fine-tuning, RAG, Agentic AI, Prompt Engineering, Data visualization, and Cloud based workflows. Proven ability to improve efficiency and accuracy in large scale data processes.

EDUCATION

UNIVERSITY OF TEXAS AT ARLINGTON

Arlington, TX

MASTER OF SCIENCE IN DATA SCIENCE

Jan 2023 - May 2024

Concepts: Machine Learning, Neural Networks, Deep Learning, Probability, Statistics, Data Visualization, Artificial Intelligence

ANNAMACHARYA UNIVERSITY

Andhra Pradesh, IN

BACHELOR OF TECHNOLOGY IN COMPUTER SCIENCE AND ENGINEERING

Jun 2017 - Jul 2021

Concepts: Python, Database Management, Big Data, Cloud Computing, Machine Learning, Web development, Data Science, Data Structures and Algorithms, Thesis: " Object Detection with Deep Learning, Computer Vision and dual stage Convolutional Neural Network"

WORK EXPERIENCE

Senior Data Scientist, SMD Technosol (Client: Thermo Fisher Scientific), Dallas, TX

Aug 2024 - Present

- Developed Python based ingestion pipelines and automated Databricks ETL workflows, reducing manual intervention and improving processing efficiency by 40%.
- Designed and deployed ML forecasting models achieving 96% accuracy, optimizing experimental planning and resource allocation.
- Built autonomous AI agents and LLM powered assistants using LangChain and Azure OpenAI, reducing manual workload by 30%.
- Implemented AI driven anomaly detection and synthetic data generation pipelines, improving data quality by 33%.

Data Research Intern, Tomorrow's Leaders Today, Frisco, TX

Jan 2024 - May 2024

- Applied 10,000+ grant records and deployed ML models to predict high potential opportunities, increasing targeting accuracy by 40%.
- Built NLP and LLM based summarization pipelines, reducing manual effort by 45%.
- Created interactive dashboards in Power BI and Tableau, improving reporting efficiency by 32%.

Research Assistant, University of Texas at Arlington, Arlington, TX

Oct 2023 - May 2024

- Developed CNN and LSTM models and GAN-generated synthetic datasets for crop yield prediction, improving model performance by 25% and achieving more than 85% accuracy.
- Integrated sensor, soil, and weather data into multimodal AI pipelines to enable crop optimization recommendations using Generative AI.
- Conducted statistical analysis, feature engineering, and documentation to support research publications and deployment.

Software Engineer, DXC Technology, Bangalore, IN

Jun 2021 - Aug 2022

- Developed automation scripts and Python validation workflows to monitor and optimize secure file transfers, improving data accuracy and reducing incident response time.
- Migrated legacy workflows to scalable ETL frameworks and built SQL and Power BI dashboards, increasing processing efficiency by 25%.
- Implemented change management processes in Jira and automated alerts with AWS, reducing deployment errors and improving workflow reliability.

KEY SKILLS

Programming & Scripting: Python, SQL, Pyspark, HTML

Machine Learning & AI: Supervised & UnSupervised Learning, Deep Learning (CNN, RNN, LSTM, Transformers, GANs), Reinforcement Learning, NLP, Computer Vision, Generative AI, Agentic AI, RAG, Prompt Engineering (Zero-shot, Few-Shot, Chain-of-Thought, ReAct, Tree-of-Thoughts), LLM Fine-Tuning (LoRA, QLoRA), A/B Testing, Statistical analysis (ANOVA, Correlation)

Frameworks & Libraries: PyTorch, TensorFlow, Keras, Scikit-Learn, XGBoost, Hugging Face, LangChain, LlamaIndex, CrewAI, LangGraph, Pandas, NumPy, Matplotlib, Seaborn, OpenCV, NLTK

Data Engineering & Cloud: ETL / ELT Pipelines, Databricks, AWS (S3, EC2, CloudWatch, SNS, SQS), Azure (ADF, Blob Storage, ML Studio), GCP (BigQuery, Vertex AI), Vector Databases (FAISS, Chroma)

Tools, Visualization, BI & Collaboration: PowerBI, Tableau, Excel, Jupyter notebook, Google Colab, Git, GitHub, Cursor AI, VS Code, Copilot, FastAPI, Jira, ServiceNow, Agile/Scrum, SDLC

RESEARCH PROJECTS & ACHIEVEMENTS

AI Log Error Explainer Agent

Built an AI agent that reads system logs and identifies errors automatically using LLMs and prompt Engineering, helping teams troubleshoot faster and make informed decisions.

Patient Readmission Predictor With AI Appeal Letter

Created a machine learning model to predict which patients might be readmitted and added an AI-powered appeal letter generator, making healthcare interventions smarter and more efficient.

Service Outage RCA Assistant

Developed an AI assistant that analyzes logs and detects anomalies to find the root cause of service outages, speeding up problem resolution and improving system reliability.

Kaggale Capstone: Intelligent Business Analytics Agent - Enterprise Agent

Built a multi-agent AI system to automate business analytics workflows, including data processing, trend detection, statistical analysis, visualization reporting, delivering faster, scalable, and actionable business insights.

Google's FooBar

Invited by Google to take part in their hidden Foobar challenge, where I completed all 5 levels. It was a great deep-dive into tough algorithmic problems and sharpened my skills in problem-solving, optimization, and creative thinking.