

ANGEL MACWAN

+91 7405-448-639 | armacwan@gmail.com | [linkedin.com/in/angel-im](https://www.linkedin.com/in/angel-im) | github.com/angelmacwan

Professional Summary

Data Scientist and Engineer with 4+ years at IBM building large-scale data pipelines, ML systems, and production LLM applications. Hands-on experience designing agentic workflows with LangChain and LangGraph, engineering RAG platforms, and deploying models end-to-end on cloud infrastructure. Passionate about building intelligent systems at the intersection of data engineering, applied ML, and generative AI.

Education

Amity University

M.Sc Data Science (CGPA: 9.96/10)

India

Sept. 2020 — Jun. 2022

Sardar Patel University

B.Sc Computer Application and Information Technology (CGPA: 8.6/10)

India

Jun. 2017 — Oct. 2020

Technical Skills

Programming Languages: Python, Go, R, Java, JavaScript, C/C++, SQL, NoSQL

ML & AI: TensorFlow, PyTorch, JAX, LangChain, LangGraph, RAG, LLMs, NLP, Agentic Systems, Prompt Engineering, Model Deployment

Data Science: Statistical Analysis, Time Series, Regression, Hypothesis Testing, A/B Testing, Predictive Modeling, Model Tuning

Data Engineering: PySpark, Hadoop, Sqoop, Databricks, Snowflake, Streaming Pipelines, Data Validation

Cloud, DevOps & Visualization: AWS, Azure, Docker, Git, FastAPI, Flask, WatsonX, Tableau, Power BI, D3.js

Experience

Data Engineer

IBM

Jun. 2022 — Present

Pune, India

- Designed distributed batch and streaming data pipelines with **Spark, Snowflake, and Databricks** processing **2TB+**, reducing end-to-end latency by **60%** at **99.9% reliability**.
- Developed and deployed **ML models** for time series forecasting, classification, and NLP tasks, achieving **92% prediction accuracy** and serving **50+ data scientists and stakeholders**.
- Engineered **RAG-based LLM applications** and prompt optimization frameworks, improving response quality by **35%**.
- Built intelligent **AI agents** using **LangChain and LangGraph**, enabling autonomous task execution and multi-step agentic workflows across enterprise use cases.
- Collaborated with the **watsonX team** to build internal AI tools using **watsonX, IBM Granite models, and Gemini API**.
- Conducted **A/B testing** and **hypothesis testing** campaigns, improving product metrics by **28%**; implemented **CI/CD pipelines** and automated data validation, reducing deployment failures by **35%**.

Data Science Intern

MindWorks Global

Jun. 2021 — Sept. 2021

Delhi, India

- Developed **predictive analytics models** using **Python and scikit-learn** on **500K+ records**, delivering insights that improved client decision-making by **25%**.
- Built interactive **data visualization dashboards** using **React and D3.js**, reducing dashboard creation from **4 hours to under 20 minutes** and cutting stakeholder reporting time by **80%**.

Software Analyst / Trainee

TCS

Aug. 2020 — Oct. 2020

Ahmedabad, India

- Architected backend APIs supporting enterprise microservices across **three production applications** and built **data quality validation frameworks** for payment processing systems using **Java, Python, and Node.js**.
- Analyzed data pipelines and developed **SQL-based analytics solutions**, identifying anomalies in payment systems and improving transaction throughput while reducing latency by **10%**.

Projects

Thallus: AI-Driven Simulation & Decision Intelligence Engine

- Architected an **AI-driven Decision Intelligence platform** that models narrative spread and behavioral outcomes using **multi-agent simulations**, enabling scenario analysis and predictive modeling of complex social dynamics.
- Designed **ML pipelines** combining **knowledge graph construction**, NLP-based ontology extraction, and **LLM-powered persona synthesis** to generate synthetic behavioral datasets from unstructured documents.
- Built **advanced analytics dashboards** with interactive visualizations, multi-agent debate workflows, scenario comparison algorithms, and statistical inference engines for extracting actionable intelligence.

Vector: Enterprise RAG & Knowledge Base Platform

- Developed an **enterprise-scale RAG (Retrieval-Augmented Generation) platform** leveraging **LLMs, embeddings, and vector databases** to enable intelligent document retrieval and conversational AI at scale.
- Implemented **advanced NLP techniques** including semantic search optimization, prompt engineering, and response ranking algorithms, improving answer relevance by **40%**.
- Built **enterprise-grade access control** using role-based permissions and multi-tenant session management, enabling secure organizational knowledge base curation and usage analytics.