ALBIN ABI THOMAS

Senior AI Engineer

(+91) 7994741413 / (+971) 504672712 | albinthomas231@gmail.com | LinkedIn | Medium

SUMMARY

AI/ML Engineer with 5 years of experience delivering production-ready AI solutions using LLMs, Generative AI, and MLOps practices. Proficient in Python, FastAPI, Docker, and cloud-native workflows. Strong background in building agentic RAG systems, deploying ML pipelines with model tracking, and enabling scalable AI deployments. Experienced in AutoML platforms, CI/CD, and cross-functional collaboration for enterprise-grade AI solutions.

SKILLS

AI/ML Expertise: Generative AI, RAG, Multimodal LLMs, MLOps, AutoML, Computer Vision, NLP, Prompt Engineering, Scikit-Learn, REST APIs, PostgreSQL, Relational Databases, Backend Development, Data Visualization, Vector Databases, Time Series Forecasting, Machine Learning, Pandas, ETL

Programming: Python, C++, JavaScript, SQL

Frameworks & Libraries: React.js, FastAPI, Flask, LangGraph, Gradio, Streamlit, Pydantic, Celery, RabbitMQ, Redis

Cloud & DevOps: IBM Cloud, Docker, Kubernetes, Git

Data/Tools: Power BI, Cognos, IBM Watsonx.data, Tableau, Web Scraping, R&D Storytelling, VS Code, Postman, Agile (Scrum)

Certifications: Agentic AI Agents (O'Reilly), IBM watsonx.ai Sales Foundation, IBM watsonx Essentials, watsonx Orchestrate Practitioner Advanced, Vault Sales Foundation

EXPERIENCE

IBM India Pvt Ltd Nov 2023 – Present
Senior AI Engineer

- Led client-facing Generative AI POCs and pilots using IBM Watsonx products, accelerating time-to-value by **70%** and driving revenue growth across Latin America and North America.
- Delivered advanced AI solutions including text detection (OpenCV + OCR), COBOL-to-Java conversion, Ansible-based infra-automation, Cognos dashboards, RAG on web-scraped data, and multimodal LLMs for entity extraction/validation.
- Engineered an automated entity extraction evaluation system leveraging LLMs as judges, reducing evaluation time and accelerating development cycles by **90%**.
- Developed a GitHub README creator, automating comprehensive project documentation and saving **5–6 hours** per project.
- Built and deployed a scalable multimodal RAG chatbot using Vector Databases with agentic behavior for contextual insights from PDF manuals, incorporating feedback loops and source-linked responses with Pydantic.
- Managed technical enablement for IBM Watsonx products, reducing enablement time by **50%** through effective scheduling and team coordination.
- Engineered end-to-end ML pipelines using AutoML frameworks and cloud-native infrastructure, exposing models via FastAPI-based REST APIs for low-latency, scalable production inference.
- Achieved global Watsonx Challenge recognition for developing a VS Code extension, **Shinigami Eyes**, for bug/vulnerability detection and unit test case generation.
- Developed a **Deepfake Detection System with over 90% accuracy**, leveraging hybrid neural networks trained on real and AI-generated data to detect manipulated content and strengthen digital trust.

Winjit Technologies Pvt Ltd May 2021 – Nov 2023
Software Developer

- Developed the flagship AutoML platform, **PredictSense**, managing end-to-end workflows for machine learning problem types.
- Enhanced the Clustering Analysis module, integrating new algorithms and building an analytics dashboard, improving cluster insights by **70%**.
- Directed client POC projects, including Crime Prediction (Survival Analysis), Security team contract expiration (Change Point Detection), and Demand Forecasting.
- Contributed to product stability and feature development across Time Series Forecasting, Predictive Modeling, EDA, Feature Engineering, NLP, and Model Deployment.
- Guided Geospatial Data Analysis Module implementation, optimized production environments with Docker, and developed use case documentation.
- Ensured timely delivery with robust code, transitioned to a monolithic repo, wrote unit tests, and contributed to peer code reviews, reducing development efforts by **40%**.

EDUCATION

Maulana Azad National Institute of Technology, Bhopal Jul 2017 – May 2021
Bachelor of Technology, Computer Science & Engineering