SURAPATI RAVI KUMAR

AI/MLOps Engineer & Full Stack Developer
Email: ravisurapati12@gmail.com │ Phone: +91-8106258635 │ Location: Hyderabad, TG │
LinkedIn: http://www.linkedin.com/in/ravikumar-surapati-b7969aa2

SUMMARY

AI/ML Specialist with 10 years of IT experience, including 5 years in machine learning, deep learning, NLP, and conversational AI of which 2 years are focused on Generative AI. Skilled in building end-to-end ML pipelines, deploying AI models in production, and developing intelligent chatbots with RASA, Dialogflow, Hugging Face, and LangChain. Experienced in MLOps (MLflow, DVC, SageMaker, Kubernetes, CI/CD) and cloud automation (AWS, Azure, Docker, Terraform). Earlier 4 years of career in full-stack development (Node.js, Angular, React, multi-cloud platforms) provide a strong backend and frontend foundation that complements AI/ML expertise.

CORE SKILLS

- **AI/ML & Deep Learning:** Python, PyTorch, TensorFlow, Scikit-learn, Hugging Face Transformers, LangChain, NLP, Generative AI, RNN/CNN/LSTM, RASA, Dialogflow
- **MLOps:** MLflow, DVC, Kubeflow, SageMaker, CI/CD (Jenkins, GitHub Actions), Docker, Kubernetes, Model Monitoring
- **Chatbots:** RASA, Dialogflow, IBM Watson Assistant, Generative AI Chatbots, Multi-channel Deployment (Slack, WhatsApp, Messenger)
- **Backend:** Python (Flask, FastAPI), Node.js, Express.js, REST APIs, GraphQL Frontend: React.js, Angular, Next.js, TypeScript, Redux
- **Cloud & DevOps:** AWS (ECS, EC2, Lambda, SageMaker), Azure, Terraform, Kubernetes, Jenkins, GitHub Actions
- **Databases:** PostgreSQL, MongoDB, MySQL, Redis
- **Vector Databases:** Pinecone, FAISS, ChromaDB

PROFESSIONAL EXPERIENCE

Full Stack Developer │ Cognizant Technology Services Pvt Ltd (Nov 2020 – Present)

- Designed and deployed end-to-end ML pipelines (data ingestion, training, evaluation, deployment) using TensorFlow, PyTorch, MLflow, and DVC, improving reproducibility by 40%.
- Built Generative AI chatbots using Hugging Face, LangChain, and Flask APIs, reducing manual support workload by 30%. - Developed an NLP-driven invoice retrieval system (transformer embeddings) reducing search time from 2 minutes to 10 seconds.
- - Deployed ML solutions on AWS SageMaker & ECS with Flask/FastAPI APIs, maintaining 99.9% uptime.
- - Automated CI/CD pipelines for ML models with Jenkins, Kubernetes, and GitHub Actions, reducing release cycles by 80%.
- - Delivered full-stack solutions with Python (Flask/FastAPI) and Node.js (1 year), integrating APIs and React-based dashboards.

Full Stack Developer | Oasys Cybernetics Pvt Ltd (Nov 2019 – Aug 2020)

- Developed Hospital Management System that reduced patient admin time by 25%.
- Designed modular Angular components enabling 40% faster feature delivery.
- Implemented Dockerized CI/CD workflows improving deployment reliability

Full Stack Developer | Memorres Digital Pvt Ltd (Mar 2019 – Oct 2019)

- Created RAGA chatbot (RASA NLU + Dialogflow) serving 5K+ users in real time.
- Improved frontend data fetching speed by 60% via optimized GraphQL + REST APIs.
- Managed multi-cloud infrastructure with Terraform, improving DR readiness by 35%.

Full Stack Developer | Abbigale Cloud Services Pvt Ltd (Aug 2015 – Feb 2019)

- Built telemedicine platform for 10K+ patients, reducing consultation wait times by 50%.
- Deployed multi-channel chatbots across Slack, WhatsApp, and Messenger, increasing engagement by 35%.
- Mentored 4 developers, improving sprint velocity by 20%.

PROJECT HIGHLIGHTS

- **MAXHUB AI ChatBot** – Reduced internal support tickets by 25% through IBM Watson Assistant chatbot integration.
- **Caterpillar Digital Marketplace** – Increased organic traffic by 30% via SEO-optimized React components.
- **SANOFI ChatBot** – Reduced finance query handling time by 40% with Generative AI and SAP integration.

EDUCATION

Bachelor of Technology in Computer Science Engineering

Vishnu Sree Institute of Technology And Sciences (JNTUH) | 2009-2013