

Yash Rane

+91 7208713112 | yash.tushar13@gmail.com | LinkedIn | GitHub | Portfolio
AI Engineer | Full Stack Developer | AWS Cloud Enthusiast

EDUCATION

Vidyavardhini's College of Engineering & Technology (VCET) <i>B.Tech in Artificial Intelligence & Data Science CGPA: 8.5</i>	2023 – 2027
Thakur College of Science and Commerce <i>Higher Secondary Certificate (XII)</i>	2022

EXPERIENCE

Acmegrade <i>Web Developer Intern</i>	Aug 2024 – Oct 2024
- Built 4+ full-stack web apps using React.js, Express.js and MongoDB, reducing page load time by ~30% via query optimization and component-level code splitting. - Designed RESTful APIs serving 10+ endpoints; implemented reusable component architecture that cut feature delivery time by 25% across Agile sprint cycles.	
Freelance Full Stack Developer <i>Self Employed</i>	2024 – Present
- Delivered 3+ client websites on MERN Stack with AWS (S3, EC2, ECS Fargate); managed media storage for files up to 50MB with secure JWT-based authentication. - Architected scalable backends handling 500+ concurrent requests; reduced API response latency by 40% through indexing and caching strategies.	

PROJECTS

Customer Care Agent <i>AWS Bedrock AgentCore, LangChain, FAISS, Titan Embeddings v2</i>	GitHub
- Built an agentic RAG pipeline over 500+ FAQ entries using Amazon Titan Embeddings v2 and FAISS; achieved ~92% retrieval accuracy on domain-specific queries. - Deployed on AWS Bedrock AgentCore with CloudWatch observability, multi-turn conversational memory and tool-calling workflows; reduced average query resolution time by 35%.	
DocMind — Multi-Document RAG Assistant <i>LangGraph, ChromaDB, HuggingFace, Groq</i>	GitHub
- Engineered a multi-PDF semantic search system with 1000-token chunking, ChromaDB vector store and Cross-Encoder reranking; improved answer relevance by ~28% over baseline BM25. - Implemented LangGraph stateful query-rewriting pipeline reducing hallucination rate by 40% across 200+ test queries; deployed via Streamlit with sub-2s end-to-end response latency.	
VeriDocs AI <i>MERN Stack, HuggingFace Transformers, MongoDB Atlas</i>	GitHub
- Built a multimodal AI verification platform detecting forged documents, manipulated images and deepfakes with 94%+ classification accuracy using HuggingFace vision-language models. - Designed Multer-based file pipelines supporting PDFs, images and videos up to 50MB; secured with JWT + RBAC; handled 1000+ verification requests in testing with zero data leakage.	

SKILLS

Languages: Python, Java, JavaScript, TypeScript, SQL **Frontend:** React.js, HTML5, CSS3, Tailwind CSS
Backend: Node.js, Express.js, REST APIs, JWT Auth **Databases:** MongoDB, PostgreSQL, MySQL, ChromaDB, FAISS
Cloud & DevOps: AWS (Lambda, EC2, ECS, ECR, S3, IAM, VPC, ALB, Amplify, Bedrock AgentCore), Docker, Kubernetes, GitHub Actions
AI / GenAI: LangChain, LangGraph, RAG, AI Agents, Vector Databases, Hugging Face, Prompt Engineering, Amazon Titan Embeddings
Core CS: Data Structures & Algorithms, System Design, OOP, Distributed Systems

CERTIFICATIONS

AWS Cloud Practitioner (GeeksforGeeks) | Generative AI & ChatGPT (GeeksforGeeks) | MERN Stack (Apna College) | Java Spring Boot (Infosys Springboard) | Google Data Analytics (Coursera)