YUVRAJ CHAUHAN

linkedin.com/in/yurajchauhan | Email: chauhanyuraj50@gmail.com |github.com/Yuraj023 |Mobile: +91 7039979533

PROFESSIONAL SUMMARY

Detail-oriented IT graduate with a strong foundation in data analysis, Python, SQL, and machine learning. Skilled at turning raw data into clear insights through cleaning, statistical analysis, and visualization. Hands-on project experience in NLP, IoT, and ML-based applications, with a structured approach to problem-solving. Eager to apply analytical skills and learn quickly in a data-driven team, contributing to better decisions through insights.

EDUCATION

Bachelor of Science in Information Technology (BSc IT) *2023 - 2026*
KES B.K. Shroff College, University of Mumbai

TECHNICAL SKILLS

Programming Languages: Python, SQL, TypeScript, JavaScript
Machine Learning & Deep Learning: Scikit-learn, PyTorch , Pandas, NumPy, Matplotlib, Seaborn
Generative AI & LLM Engineering: RAG Pipelines, Prompt Engineering , LLM API Integration, Embeddings, Semantic Search,
LLM Frameworks & Tools: LangChain, LangGraph, OpenRouter API , OpenAI SDK, Ollama, Hugging Face Transformers
Databases & Vector Stores: PostgreSQL (Supabase) , Chroma, MySQL
Backend & Application Development: FastAPI, REST APIs, Streamlit, React
Cloud & DevOps: AWS (EC2, IAM), Google Cloud Platform (GCP) , Docker, Git, GitHub

PROJECTS

DataGraph - OKF-Based Enterprise AI Copilot *Sep 2026*
Technologies: Python, Streamlit, OpenRouter API, OpenAI SDK, OKF, Markdown, Regex

- Developed an OKF-based AI copilot that auto-generates SQL queries, CLI commands, and security-audit scripts from linked Markdown knowledge
- Designed graph-based retrieval across business rules, metrics, and operational runbooks without a vector database
- Implemented context compression, JSON recovery, and retry logic to improve prompt efficiency and API reliability
  GitHub

Character-Level LSTM Language Model from Scratch *Aug 2026*
Technologies: Python, NumPy, Deep Learning, LSTM, NLP

- Implemented a character-level LSTM language model from scratch using NumPy to learn sequential patterns in text
- Processed 204,800 characters using a sequence length of 50 and a hidden size of 128
- Exported model weights and vocabulary for reuse and inference
  Github

IssueSense AI - Intelligent Issue Categorization System *Jan 2026 - Feb 2026*
Technologies: Python, Scikit-Learn, TF-IDF, Logistic Regression, NLP, pickle

- Engineered a multi-class NLP classification system (10 categories) achieving 94.38% accuracy and a 94.4% macro F1 score on 10,000+ support tickets
- Automated support ticket categorization to streamline manual triage workflows
- Serialized the model using pickle and built a standalone prediction script for seamless ticketing system integration
  GitHub | Live Demo

Kiki AI - RAG-Based Intelligent Chatbot *Sep 2025 - Oct 2025*
Technologies: TypeScript, OpenRouter API (DeepSeek V3), RAG, Supabase

- Built a companion-style chatbot using TypeScript and RAG to deliver contextually relevant responses
- Integrated Supabase to store conversation history and maintain context across multi-turn sessions
- Implemented a RAG pipeline to retrieve relevant context and inject it into LLM prompts for context-aware response
  GitHub | Live Demo

CERTIFICATIONS

- Generative ai for content creation – ibm skillsbuild (prompt engineering, responsible ai)
- Data analytics with excel – value-added course
- Tata — genai powered data analytics job simulation — forage
- Generative ai, chatgpt and prompt engineering – value-added course
- Web design and development – value-added course