

ATHARVA SANDEEP RAUT

Phone: **+91 9881526584** | E-mail: atharvaraut.asr@gmail.com

GitHub: Yousician-01 | LinkedIn: Atharva Raut

Website: <http://atharva-raut.vercel.app/>

Address: Karjat, Maharashtra, India

ABOUT ME

Machine Learning & AI Systems Engineer focused on building production-ready intelligent systems combining machine learning, LLMs, graph reasoning, and backend engineering. Interested in AI infrastructure, reasoning systems, cybersecurity, and scalable ML applications from research through deployment.

AI/ML, DATA SCIENCE & SYSTEM SKILLS

- **Languages:** Python, SQL, Typescript, C++
- **Machine Learning & AI:** PyTorch, Scikit-Learn, Classification, Regression, Clustering, Anomaly Detection, NLP (TF-IDF, Logistic Regression, NER), Transformers (DistilBERT, HuggingFace), Computer Vision (OpenCV, MTCNN, Inception ResNet V1)
- **Data Science & Analytics:** Numpy, Pandas, EDA, Feature Engineering, Statistical Analysis, Data Cleaning & Transformation, Hypothesis Testing, Data Visualization, Model Evaluation (Precision, Recall, F1-score, ROC-AUC, Confusion Matrix)
- **Backend & Systems:** FastAPI, Flask, Django, REST APIs, Async Programming, Scalable System Design
- **AI Frameworks:** LangChain, MLflow, Prompt Engineering, n8n
- **Deployment & DevOps:** Docker, GitHub Actions, CI/CD, Git, Full-stack Deployment Next.js, CORS, Environment Configuration, API Integration

PROJECTS

AI SOC Investigation Platform (Ongoing)

- Designed an AI-powered investigation platform that reconstructs cyberattacks from raw security telemetry.
- Built deterministic reasoning engines for evidence extraction, behavior correlation, MITRE ATT&CK mapping, and graph-based campaign reconstruction.
- Integrated LLMs to generate explainable investigation reports grounded in deterministic evidence.

Tech Stack: Python, FastAPI, Transformers, Graph Reasoning, Cybersecurity, LLMs

LLM-powered Full Stack Application (FastAPI + Next.js)

- Built a FastAPI + Next.js application integrating Hugging Face transformer models.
- Developed asynchronous REST APIs for AI inference with production-oriented architecture.
- Implemented responsive frontend for real-time AI interaction.

Tech Stack: FastAPI, Next.js, HuggingFace, Python

AI-based Home Security System (Computer Vision + Embedded Systems)

- Developed an identity recognition pipeline using OpenCV, MTCNN, and Inception ResNet V1.
- Integrated Raspberry Pi hardware for automated access control.
- Optimized inference pipeline for real-time face detection and recognition.

Tech Stack: Python, OpenCV, Raspberry Pi, Deep Learning

Data Inspector – Dataset Health & EDA Tool

- Implemented schema inference, missing value analysis, outlier detection, and correlation visualization.
- Automated dataset health assessment before machine learning training.
- Improved data quality inspection using modular analytical workflows.

Credit Card Fraud Detection

- Developed a fraud detection model for highly imbalanced financial transaction data using feature engineering and model optimization.
- Evaluated multiple classifiers using Precision, Recall, ROC-AUC, and confusion matrix analysis.

Stock Direction Indicator

- Built a Random Forest-based stock direction prediction system with confidence-weighted trade signals.
- Developed a Streamlit dashboard and packaged the application using Docker with a CI/CD workflow.

EDUCATION

BITS Pilani 2026	2026 – Present
Master of Science Data Science & Artificial Intelligence <i>(Pursuing)</i>	
Vishlesan i-Hub, IIT Patna & Masai School	10/2025 - Present
Artificial Intelligence & Machine Learning Program <i>(In Progress)</i> AI	
• An intensive industry-oriented program covering ML, DL, NLP, and applied AI development.	
B.N. Bandodkar College of Science (Autonomous)	06/2021 - 03/2024
Bachelor of Science Physics	CGPA: 8.75/10
• Relevant Coursework: Python, C++ and History and Architecture of Computers	

RESEARCH EXPERIENCE

Research on A Hybrid Two-Tiered Approach for Fake News Detection Using Machine Learning and Natural Language Processing Independent	02/2025
• Designed a dual-stage misinformation detection framework combining TF-IDF + Logistic Regression with DistilBERT contextual analysis. • Achieved 98.85% classification accuracy while integrating Named Entity Recognition and real-time fact verification. • Research paper accepted for publication in a Springer journal.	
Developed a forensic/ cybersecurity device	10/2024
• Developed and a device named Forensic USB that provides protection against USB Rubber Duckies.	

INDUSTRY JOB SIMULATIONS

Deloitte - Data Analytics Job Simulation	
• Solved structured data analytics task in forensic and compliance scenarios. • Interpreted analytical outputs to support business risk and investigation decisions.	
Forge - Data Visualisation: Empowering Business with Effective Insights	
• Worked on end-to-end visualization tasks from problem framing to dashboard creation. • Selected appropriate charts and communicated insights for decision making stakeholders. • Communicated insights clearly using data driven reasoning.	

CERTIFICATIONS

- Complete Data Science, Machine Learning, Deep Learning & NLP Bootcamp (Udemy).
- Mathematics - Basic to Advanced for Data Science and GenAI (Udemy).
- Near-Earth Object Analysis - International Astronomical Search Collaboration (NASA, Pan-STARRS)

EXTRACURRICULAR ACTIVITIES

- Presenter at Creative Ideas & Innovations in Action (CiiA-4) Exhibition (2025).