Yashraj Shrivastava

+91-8839664523 yashraj170803@gmail.com linkedin.com/yashraj github.com/yashraj

Education

Indian Institute of Technology, Madras   Expected May 2027
B.S. in Data Science And Application   *Chennai, Tamil Nadu*

Projects

AstroBot: AI Chatbot for Space & Weather Data | *LangChain, RAG, LLaMA-3.1, FAISS, Flask, Web Crawling*
- Architected a domain-specific **Generative AI chatbot** for **ISRO's MOSDAC data**, enabling natural language interaction on satellites, missions, and real-time weather.
- Implemented **Retrieval-Augmented Generation (RAG)** with FAISS, integrating **1000+ crawled ISRO documents** into a knowledge base enhanced by **LLaMA-3.1 (Groq API)**.
- Augmented responses with **dynamic weather intelligence**, improving contextual accuracy by **35%**.
- Shipped advanced features such as **speech-to-text input**, **persistent chat history**, and one-click **PDF report export**.
- Deployed a **production-ready Flask application** with SQLite, supporting multi-source query pipelines across ISRO, NASA, and live weather APIs.
- Enabled seamless access for researchers and students, making complex space and weather data easily interpretable.

TruthTell: Real-time Fake News Detector | *Flask, PassiveAggressiveClassifier, NLP, Deep Learning*
- Developed an **AI-powered misinformation detection engine** achieving an **F1-score of 87.6%** on 5000+ articles.
- Optimized inference pipeline, reducing latency from **8s → 4.5s** (**65% improvement**).
- Engineered scalable Flask deployment with live text ingestion and ongoing extension to **multimodal verification** (audio, image, video).
- Analyzed common misinformation patterns to improve model robustness and real-time accuracy.

Video Person Tracker | *Computer Vision, YOLOv8, DeepFace, OpenCV*
- Engineered a **real-time multi-person facial recognition system** with **92% accuracy** leveraging YOLOv8 and DeepFace.
- Optimized video analytics with **temporal frame sampling** and GPU acceleration, yielding a **40% throughput gain**.
- Mitigated false positives by **35%** using a custom temporal clustering algorithm.
- Delivered **3.2s processing time per video minute** via multi-threading and model quantization.
- Demonstrated practical applications in security surveillance and attendance management systems.

Technical Skills

Programming Languages: Python, C++, SQL
Generative AI & LLMs: Prompt Engineering, LangChain, RAG (Retrieval-Augmented Generation), OpenAI & LLaMA APIs, Hugging Face Transformers, FAISS, ChromaDB, Vector Databases
Natural Language Processing (NLP): Text Preprocessing, Summarization, Translation, Sentiment Analysis, Named Entity Recognition (NER)
Deep Learning: Neural Networks, CNNs, RNNs, TensorFlow, PyTorch, Transfer Learning
Computer Vision: YOLOv8, DeepFace, OpenCV, Object Detection & Tracking
Data & Deployment: Flask, FastAPI, REST APIs, Streamlit, SQLite, MySQL, Model Quantization, Multi-threading, Docker (if applicable)
Libraries & Tools: NumPy, Pandas, Scikit-learn, Git, GitHub

Certifications

- **Pursuing a 4-year certification** from PW Institute of Innovation
- **Data Analytics and Visualization Job Simulation** - Accenture (Jul 2024): 🌐
- **Analyzing and Visualizing Data with Microsoft Power BI** - Great Learning (Jan 2024): 🌐
- **Pandas and NumPy Libraries** - Infosys Springboard (Feb 2024, valid till Dec 2034): 🌐
- **MySQL 8** - Great Learning (Jan 2024): 🌐