

Sambhav Surana

AI/ML Developer | Python Developer



Third-year student specializing in Artificial Intelligence. I have built custom AI models, intelligent agents, and real-time face recognition systems from scratch. Skilled in Python, PyTorch, and full-stack development. I am passionate about turning complex AI concepts into working software and am looking for a challenging role in engineering.

📍 India 📞 +917999468896 🎙️ sambhav242005@gmail.com 🌐 <https://sambhav-surana.online>

Experience

Abhyudaya Coding Club

Club Member

Shri Vaishnav Vidyapeeth Vishwavidyalaya

2024 - Present

- The development of AI-based projects including Time Table Generation system using constraint satisfaction algorithms
- Organized workshops on Python, NumPy, Pandas, and Scikit-learn for club members
- Collaborated with team members on model optimization and performance enhancement

Certifications

Hackathons & Competitions

Hackwave 2.0

Aug 2025

Participated in a 36-hour coding hackathon organized by Echelon Dev Society at CDGI, Indore. Gained hands-on experience in rapid prototyping, teamwork, and problem-solving.

🔗 <https://hackwave-site.vercel.app/>

Introduction to Machine Learning

NPTEL Online Certification –IIT Kharagpur

Jul–Sep 2024

- Achieved Elite Certificate with a consolidated score of 66%
- Completed an **8-week course** covering fundamental machine learning concepts and assignments

Technical Skills

Programming Languages



Python, JavaScript, SQL, Bash

ML Libraries & Frameworks



NumPy, Pandas, Scikit-learn, TensorFlow, PyTorch, Matplotlib

Computer Vision



OpenCV, Image Processing, Data Augmentation, YOLO, Face Recognition, Object Detection

Deep Learning



CNN, Neural Networks, Model Training, Transfer Learning, Feature Extraction

NLP & Generative AI



LangChain, OpenAI API, Ollama, Text Processing, Query Decomposition, LLM Integration

Deployment & Tools



REST APIs, Git, GitHub, Model Deployment, Gradio, WebSockets

Databases



MySQL, MongoDB, ChromaDB

Web Technologies



React.js, Next.js, HTML/CSS, Tailwind CSS

Development Tools



VS Code, PyCharm, Jupyter Notebook, Git

Operating Systems



Windows, Linux

Education

Shri Vaishnav Vidyap...



CSE (AI-IBM) III Year

7.2 CGPA

B.Tech.

2023 - Present

Projects

Multi-Agents - Automated Product Development

An autonomous agent system that converts raw ideas into technical product specifications.

- **Autonomous Agent Pipeline:** Orchestrated a 7-stage **LangChain** workflow simulating a cross-functional team (Product Manager, Engineer, Risk Analyst) to autonomously generate comprehensive project blueprints.
- **Role-Based Logic:** Implemented specialized agents including a "Clarifier" to refine user intent and a "Risk Agent" to identify security pitfalls before architectural proposal.
- **Scalable Infrastructure:** Deployed the **FastAPI** backend and **Next.js** frontend using **Docker** and **Nginx** to ensure robust performance and ease of deployment.

LangChain, Multi-Agent Systems, FastAPI, Docker, Next.js, AI Orchestration

🔗 <https://github.com/Sambhav242005/Multi-Agents>

Savant - Custom Universal Transformer

A parameter-efficient Recurrent Language Model built from scratch in PyTorch.

Nov 2025 - Dec 2025

- **Custom Architecture:** Engineered a Universal Transformer with a 12-step recurrent block and **Liquid Gating mechanisms** to maximize reasoning density while minimizing parameter count.
- **Logic-Heavy Training:** Implemented a "Savant/Genius" data mixing strategy (80/20 ratio), interleaving general corpus with high-intensity logic problems to enhance model reasoning capabilities.
- **Resource Optimization:** Designed a custom training loop with **bfloat16 mixed precision**, gradient accumulation, and 8-bit optimizes to ensure stability on limited VRAM (10GB) hardware.

PyTorch, Universal Transformer, LLM Architecture, Liquid Neural Networks, Model Optimization

🔗 <https://github.com/Sambhav242005/Project-Savant>

Languages

English



Hindi



German



Real-Time AI Face Recognition System

High-throughput face tracking system using YOLOv11 and Vector Database retrieval.

March 2023-May 2025

- **High-Performance Architecture:** Engineered a **multi-process** system (separating detection, encoding, and rendering threads) to maximize FPS and reduce latency on CPU/GPU streams.
- **Vector Search Integration:** Implemented **ChromaDB** to store and retrieve face embeddings, enabling scalable O(1) identification of individuals from a large database.
- **Advanced Tracking:** Integrated **YOLOv11-face** for state-of-the-art detection accuracy combined with object tracking algorithms to maintain persistent IDs across video frames.

YOLOv11, ChromaDB, Computer Vision, Multi-Processing, OpenCV, Real-time Systems

🔗 <https://github.com/Sambhav242005/AI-Face-Detect...>

Prime-Paradox (CampusCare)

Intelligent medical ecosystem connecting patients and doctors via Multi-Agent AI.

Nov 2025

- **Multi-Agent AI:** Architected an "**LLM Council**" using **LangGraph** and **OpenRouter** to analyze symptoms and generate multi-perspective diet and exercise plans.
- **Full-Stack Architecture:** Developed a reactive frontend using **Next.js 15** and **Shadcn UI**, coupled with a high-performance **FastAPI** backend and Supabase for real-time data synchronization.
- **Real-Time Monitoring:** Integrated device hubs for tracking user vitals and implemented secure messaging channels for direct doctor-patient communication.

Next.js 15, FastAPI, LangGraph, AI Agents, Supabase, Full Stack Development

🔗 <https://github.com/Voidhacks7-0/Prime-Paradox>