Sanatan Khemariya

7898009986 — sanatankhemariya@gmail.com — linkedin.com/in/sanatank — github.com/sanatren

Education

Jaypee University of Engineering and Technology **August 2022 - July 2026**
Bachelor of Technology in Computer Science and Engineering *Guna, India*

Key Courses: Operating Systems, Computer Networks, Data Structures, Algorithms, Machine Learning, Databases

Technical Skills

Languages Python, C++, Swift
Frameworks FastAPI, LangChain
Databases MongoDB, Supabase (PostgreSQL), MySQL
Tools Streamlit, Git, Github, Langflow, Xcode
Libraries PyTorch, TensorFlow, Scikit-learn, OpenCV, NumPy, Pandas
OS MacOS, Windows

Certifications

- Kaggle - Pandas for Data Analysis
- Tata Technologies - Introduction to Generative AI

Projects

Azazel - Supercharging Legacy Models **Python, OpenAI, FAISS, LangChain, Streamlit**
— Project Link

- Optimized legacy models with **multimodal AI capabilities**, driving a **3× boost in intelligence** and achieving a **25% increase in response accuracy and efficiency**.
- **Reduced API costs by 30%** using GPT-4o-mini for multimodal capabilities and efficient prompt engineering
- Optimized **LangChain RAG** for context-aware responses; integrated GPT-4o-mini vision for image analysis and code debugging
- Developed comprehensive Streamlit UI with real-time web search, multilingual support, and **speech-to-text** capabilities

Easy-Notes - Effortless Writing **Python, FastAPI, TensorFlow, Hugging Face, Supabase**
— Project Link

- Trained an LSTM-based next-word prediction model achieving **66% accuracy** for enhanced writing experience
- Developed and fine-tuned a Transformer based English to Hindi translation model reaching **94.6% translation accuracy**
- Optimized PostgreSQL database architecture with Supabase, **reducing query response times by 25%**
- Engineered RESTful API endpoints with FastAPI, Uvicorn & Jinja2 templates for seamless frontend-backend integration
- Implemented low-latency ML model serving with **average prediction time under 100ms** for real-time text suggestions

Ecogenie - Smart Recycling **Python, Streamlit, Pillow, Geocoder, Google Generative AI**
— Project Link

- Created an AI-powered waste classification system using Google Gemini-1.5-flash with **93% accuracy** in recycling identification
- Optimized image processing pipeline to analyze scrap items in **under 15 seconds per image**, with bulk processing of 10+ items per minute
- Engineered structured prompt system for AI analysis, ensuring actionable safety tips and environmental impact insights
- Built responsive Streamlit UI with image upload functionality and geolocation-based recycling center recommendations

ChurnMonitor **Python, TensorFlow, Scikit-learn, Pandas, Streamlit**
— Project Link

- Engineered machine learning model achieving **87% accuracy** in predicting customer churn for the banking sector
- Trained and validated model on **50,000+ customer records** with comprehensive feature engineering
- Created interactive Streamlit dashboard delivering real-time insights and visualization of key churn prediction factors
- Developed automated data preprocessing pipeline reducing data preparation time by **20%**

Achievements

- Solved **300+ LeetCode problems**, demonstrating strong algorithmic thinking and data structure proficiency
- Led a team to the **National Tata Innovant (GenAI Hackathon)**, being the only selected team from Madhya Pradesh out of **2600+ competing teams**