

# PRABHUDAYAL VAISHNAV

 p09m21@gmail.com

 Portfolio Website

 LeetCode Profile

 GitHub Profile

 LinkedIn Profile

## WORK EXPERIENCE

### Research Internship: ESIEA Paris, France (Remote)

January 2025 - June 2025

*Facial Emotion Recognition Using Advanced YOLO Architectures*



- Achieved **93.4% mAP@0.5** using dual YOLO frameworks (YOLOv12 + YOLOv11) for **22-class** facial emotion recognition on CFEE dataset under Prof. Khadija SLIMANI.
- Implemented **automated annotation pipeline** reducing manual effort by **90%** and fine-tuned models via transfer learning from MS COCO pre-trained weights.

## PROJECTS

### Grand Plaza: Voice AI Hotel Concierge System

FastAPI, LangGraph, PipeCat, ChromaDB

*Built a full-stack voice AI concierge achieving 88% accuracy in menu query understanding*



- Architected a **multi-strategy RAG system** achieving **88%** query accuracy, **outperforming naive RAG** by 4.8% through semantic search, reranking, and Chain-of-Thought reasoning.
- Implemented **LangGraph-based agent workflows** with **5+ specialized nodes**, reducing conversation errors by 15% through **structured intent classification** and validation.

### Career Scout: Voice AI-Powered Job Search Assistant

Next.js, React, Supabase, TypeScript, MCP Server

*Developed a full-stack Voice AI Career Coach for smart job search*



- Architected a custom MCP server with 3 tools: *Intelligent Job Search, Company Analysis, and Web Search*.
- Integrated custom scoring logic to match user skills with confidence percentages.
- Built a responsive, Supabase-authenticated frontend with React hooks and Tailwind CSS, reducing voice interaction latency by 40%.

### • The AI Scientist: Sakana AI

Python, LLMs, Transformers, Git

*Contributed to Sakana AI's Open Source project to automate scientific research.*



- Added 8 open-source language models, achieving a 75% reduction in generation costs.
- Adapted the project for T4 GPU support, reducing reliance on high-end GPUs like H100/A100.
- Co-authored a paper: "*Exploring Style Transfer with Small Character-Level Transformers*" using Qwen2.5(72B).

### • Development of Python Module: XS

Python, Ngrok, API Creation

*Published a pip package to expose inference endpoints for open-source LLMs.*



- Released a Python package to create accessible endpoints for over 99% of Hugging Face-hosted LLMs.
- Leveraged Ngrok to tunnel requests from localhost to the cloud securely.

## TECHNICAL SKILLS

### • Programming Languages:

Python, Next.js, React, JavaScript, C, C++, SQL, Matlab, R

### • AI and LLM Tools/Frameworks:

LangGraph, LangChain, LlamaIndex, Ollama, HuggingFace Transformers, PyTorch, TensorFlow, PipeCat

### • Data Analysis and Visualization Tools:

Pandas, Numpy, Seaborn, Power BI

### • Miscellaneous Skills:

Linux, MacOS, Competitive Programming

### • Soft Skills:

Leadership, Communication, Problem Solving

## RELEVANT COURSEWORK

### • Core Computer Science:

Object Oriented Programming, Database Management System, Computer Architecture Organization, Computer Network, Operating System, Data Structures and Algorithms

### • Artificial Intelligence and Data Science:

Artificial Intelligence, Data Visualization, Pattern Recognition and Machine Learning, Natural Language Processing

### • Mathematics and Theoretical Foundations:

Theory of Computation, Cryptography and Network Security

## ACHIEVEMENTS

### Indian Institute of Management (Data Analyst) Hackathon Winner, 2024

2024

*8 hour long Hackathon, organized by InFED at Nagpur on 12 November 2024*



- Won 1st place by presenting a solution to the Startup's Problem: **AI Product Comparison**.
- Implemented the Deep-Reserach & generated a report on **AI Product Comparison (22-page report)** using a Python-based AI Agent with Tavily Search, significantly reducing manual effort and time by 90%.

## ACHIEVEMENTS

---

### HackLLM 2025 @ IIIT Delhi (Esysa'25) – 3rd Place Winner

August 2025

24-hour Large Language Model Hackathon with 250+ participants, 1.64L prize pool



- Built production-ready medical AI summarization system generating dual-perspective outputs (patient + clinician) from scratch in 24 hours without external APIs, securing top-3 finish among 80+ teams.
- Architected modular Python system with BART-large-CNN fine-tuned on medical Q&A data, achieving perspective-aware content categorization (Information/Suggestion/Experience/Cause) with regex pattern matching.

## EDUCATION

---

### Bachelors of Technology (Hons.) CSE Data Science

2022-Present

(6th Semester) Chhattisgarh Swami Vivekanand Technical University, Bhilai, 491107

CGPA : 7.2/10

- Research Internship at ESIEA Paris, France (Remote) under the project "Complex Human Emotion detection with Ultralytics YOLO11" - Completed June 2025.