Sharvani Jadhav

+91-9552365775 | sharvani.jadhav13@gmail.com | the-7g-girl | Website | LinkedIn

Education

Indian Institute of Technology, Kanpur 2021 – 2025
Bachelors in **Mathematics and Computing** CGPA: 7.5/10
Minor in **Machine Learning Applications**

Scholastic Achievements

- Qualified **Indian National Maths Olympiad** consecutively **3 times** ranking among the **top 900** students of India
- Secured **AIR 2488** in **JEE Advanced'21** and AIR 3985 in JEE Mains'21 - top 0.1% among 1.5 million candidates
- Selected among **top 25 students globally** for the Newcastle University International Summer Research Program
- **Patent: Modular Device for Precise Irrigation** – developed a cost-effective sprinkler system with moisture sensors

Work Experience

Data Scientist | *VMock India Pvt Ltd* Jun'25 – Present
- Built **multilingual support** to skills pipeline using **DeepL** for **9 languages**, boosting global reach by 30k+ users
- Built LinkedIn optimizer using **LangChain GPT integration**, **Pydantic** schema validation & **LangFuse** tracing
- Built an ensemble model of **Regression, XGBoost, SVM** with **LIME** for Wharton MBA resume screening
- Built **RAG pipeline** retrieving profile-specific questions with **GenAI** rephrasing for personalized mock interview

International Summer Research Intern | *Newcatle University, UK* Jun'24 – Aug'24
Project: Bayesian Modelling of Road Traffic Collisions - Mentor: Prof. Lee Fawcett
- Developed **Accident Prediction Model** to find best govt strategy to reduce road-accidents by **40%** at **50% cost**
- Applied **AIC-based model selection** to find **7** parameters and applied **negative binomial regression** on data
- Implemented **Metropolis-Hastings MCMC algorithm** for **10,000 iterations** to fine-tune model's parameters
- Evaluated priors like **Weibull, Lognormal** using **DIC** to identify best fit for causality data from 57 sites

Generative AI Intern | *artcube.ai* May'24 – Jun'24
- Developed **AI-based video advertisement generator** to incorporate brand's items into videos using **VAEs**
- Implemented **stable diffusion models** in ComfyUI on a **200GB VM** deployed on **Microsoft Azure**
- **Upscaled** low-resolution images and videos using models like **SUPIR, LDSR** etc achieving **4.5x improvement**

Projects

BM25+FAISS Hybrid RAG with ColBERT Re-ranking | *Self Project* Feb'25 – Mar'25
- Designed hybrid retrieval pipeline combining **BM25 lexical-search** & **FAISS vector-search** boosting 14% recall
- Applied **ColBERT token-level re-ranking** on top-100 candidates, achieving +15% MRR over baseline retrieval
- **Deployed Flask API** with <1s latency, integrating GPT context answering with NLP fallback for robustness

MPI Parallel Computing in Distributed Systems | *Prof. Preeti Malakar, CSE, IITK* Feb'25 – Mar'25
- Engineered **high-performance parallel I/O** pipeline to read & process large-scale data **reducing time by 10x**
- Implemented **3D Halo Exchange** to compute min/max cells via **6-neighbor** comparisons & edge-case handling

Probabilistic Deep Learning via Prioritized Training | *Prof. Piyush Rai, CSE, IITK* Jan'25 – Apr'25
- Implemented **RHO-LOSS** to accelerate DNN training **18x**, improving accuracy by **3.56%** on CIFAR-10 dataset
- Converted DNNs to **Bayesian Neural Networks** by normalization layer retraining for **uncertainty estimation**

Nightlight Analysis of Factories | *Mentors: Dr Hari, Frankfurt School & Dr Rangan, IIMB* Jan'23 – Jun'23
- Used **Google Earth Engine** to analyze **NASA nightlight data**, identifying **8000** factories & their annual waste
- Determined a **7.4% production rise** in December and a **67% correlation** between nightlight & waste disposed

Technical Skills

Languages: Python, R, C++, SQL, JavaScript, MATLAB, LaTeX
Developer Tools: Git, Docker, Flask, SQLAlchemy, Celery, AWS, Pydantic, Kubernetes, FastAPI, RESTful APIs
Libraries: PyTorch, TensorFlow, Hugging Face, LangChain, Scikit-Learn, SpaCy, NLTK, Matplotlib, OpenCV
ML: LLMs, GenAI, NLP, Computer Vision, Deep Learning, Transformers, Fine-tuning, RAG, Prompt Engineering
Data Science: Model Deployment, MLOps, Hyperparameter Optimization, Statistical Analysis, Feature Engineering

Relevant Courses

Natural Language Processing	Advanced Topics in ML	Probabilistic ML	Introduction to ML
Image Processing	Parallel Computing	Data Structures & Algorithms	Bayesian Statistics
Probability and Statistics	Fundamentals of Computing	Linear Algebra	Numerical Analysis & Computing