

Soham Pahari

Linkedin: <https://www.linkedin.com/in/sohampahari/>

Github: <https://github.com/suhanpahari>

Email : paharisoham@gmail.com

Mobile : +91-90645-20673

EDUCATION

University of Petroleum and Energy Studies
Bachelor of Computer Science (Data Science major)

Dehradun, India
Aug 2022 - May 2026

SKILLS SUMMARY

Programming & Tools : Python, Cuda, Java, C++, SQL, Git

Machine Learning & Libraries : Scikit-learn, TensorFlow, PyTorch, OpenCV, NLTK, Hugging Face

Frameworks & Databases : Flask, LangChain, LlamaIndex, Chrom

INTERNSHIP EXPERIENCE

Bahas Pvt Ltd

Remote

ML Development Intern

May 2024 - July 2024

- Regional Language Emotion Classification: Processed **12k+ Bengali text samples**, applied cleaning, tokenization, and TF-IDF vectorization. Trained Naive Bayes, Logistic Regression, Random Forest, SVM, and fine-tuned BERT. Achieved best results with **BERT: 66.3% test accuracy, F1 = 0.6592**, outperforming all baselines by 17%.
- Deployment: Built a Streamlit app for real-time predictions with **under 1s inference latency**, integrated 5 models with selectable interface, and improved usability for non-technical users.

NIT Warangal, Dept. of Artificial Intelligence

Warangal, India

Research Intern & Team Lead

May 2025 - June 2025

- Drug Discovery Framework: Led a team of **4 interns** in developing a custom encoder-decoder with gating + molecule-aware embeddings. Trained on **1.13M drug-indication pairs**, achieving **BLEU 0.2543, MACCS 0.4496**, and **Text2Mol 0.6250**, delivering **20-30% improvement** over SOTA baselines.
- High-Performance Training: Optimized large-scale experiments on **NVIDIA A100 DGX system** using CUDA and mixed-precision training, reducing training time by **35%**.
- Research Outcome: Authored a paper based on this work, currently **submitted to a peer-reviewed journal**.
- Funded Research Inclusion: Individually selected for a **funded medical image processing project** at NITW, in recognition of research expertise.

PROJECTS

AMoE-ViT: Adaptive Mixture of Orthogonal Experts | Python, PyTorch

[Github]

- Designed an **Adaptive Mixture of Orthogonal Experts Vision Transformer** with a dual-stage pipeline combining **image pretext learning** and **video context modeling** for unsupervised leaf disease diagnosis.
- Developed a **Self-Supervised Attention Fusion Module** that improves cross-frame reasoning, achieving strong **generalization and robustness** across unseen plant species and varying lighting conditions.

TriadNet: Multimodal Architecture for Personality Assessment | Python, PyTorch

[Github]

- Designed a **multimodal deep learning system** for Big Five personality trait prediction, integrating **Dynamic Face Graph Network**, **Prosody-Aware Capsule Network**, and **Contextual Sentiment Flow** for video, audio, and text.
- Achieved **97.8% clustering accuracy** and **96.9% mean trait prediction accuracy** on the ChaLearn LAP dataset, outperforming unimodal and prior fusion-based baselines.

ACHIEVEMENTS & CERTIFICATIONS

- Received a research grant named **SHODH** from our university's R&D department for an agricultural IoT project titled *Agirveda*. [Approval Letter]
- Winner of the Smart India Hackathon 2025 Internal Challenge** at the university level, advancing to the national round.
- Presented a research paper at **COMSYS 2025**, Warsaw University of Technology, Poland.
[Certificate]
- Presented a research paper at **KEDAR 2025** on *Towards Intelligent Transportation: YOLO Guided Electric Vehicle Recognition System in India*. [Certificate]
- Presented a research paper at **CIACON 2025** on automated umpiring for improved score analysis. [Certificate]
- Completed **Decision Making and Reinforcement Learning** by Columbia University (Coursera, April 2025).
[Certificate]