

PRANCHAL AGARWAL

(281) 673-5078 | pranchal.agarwal@gmail.com | linkedin.com/in/pranchal15 | github.com/Pranchal15

Education

University of Massachusetts Amherst, MS in Computer Science

Aug 2023 – May 2025

Coursework: Advanced NLP, Data Science Algorithms, Statistics for ML

GPA: 3.9/4.0

Jawaharlal Nehru Technological University, BS in Computer Science

Aug 2019 – May 2023

Coursework: ML Foundations, Neural Networks, Databases, Software Engineering

Technical Skills

- Languages: **Python**, SQL, R, C/C++, HTML, CSS, JavaScript, TypeScript
- Framework & Libraries: **TensorFlow**, **PyTorch**, HuggingFace, Scikit-Learn, XGBoost, NumPy, Pandas, Keras, React.js
- Deployment & Infra: Docker, FastAPI, **AWS (S3, EC2)**, **GCP**, Git, Github, HPC Clusters - Unity, Slurm

Experience

Adobe Research (Open sourced collaboration submitted to NeurIPS 2025)

Jan 2025 – Present

Machine Learning Researcher

Remote, USA

- Engineered a simplified and efficient GenAI evaluation framework to align LLM-as-a-Judge with human judgments.
- Pioneered a **7x** computationally efficient alternative to supervised fine-tuning, trained on less human feedback.
- Improved accuracy by **~20%** on pairwise preference dataset, compared to SOTA fine-tuned LLM judge - Prometheus.
- Reduced MSE by **40%** in low-data scenarios in absolute scoring via judge model calibration and self-correction.

SS&C Intralinks - FinTech

Oct 2024 – Dec 2024

Machine Learning Engineer Intern

Waltham, MA

- Architected an image to JSON pipeline using InternVL2, achieving **~60%** higher performance than OCR.
- Benchmarked multiple **VLMs** on complex financial document understanding tasks resulting in **92%** accuracy.
- Designed multimodal ingestion pipelines that can automate data flows with **5x faster inference** using LMDeploy.

Manning CICS - Prof. Chuang Gan

May 2024 – Aug 2024

Machine Learning Research Assistant

Amherst, MA

- Created a multimodal 3D asset generation pipeline using **GPT-4o** with advanced textual and spatial prompting.
- Improved asset recognition accuracy by **30%** via **chain-of-thought** prompting and multi-step reasoning.

GlobalSmart IT

May 2022 – June 2023

Machine Learning Intern

Remote, USA

- Built an end-to-end **voice assistant**, using **OpenAI API** for dynamic, context-aware conversational intelligence.
- Designed and developed an internal FAQ answering tool from meeting transcriptions, sales playbooks, emails etc, leveraging multiple models systems including NER, automatic tagging and QA chatbot.

Projects

RAG Question Answering System

- Designed and deployed a RAG QA system leveraging **vLLM** for low-latency, high-throughput inference.
- Leveraged **FAISS** for vector search and Mistral-7B-Instruct-v0.3 as the generation backend.
- Instituted fallback logic to reduce hallucinations by **~40%** in low-context scenarios.

Multi-Hop Tree

- Applied the Tree of Clarifications approach to a multi-hop QA dataset, analyzing 12.7M+ web snippets across 34,689 complex questions to enhance reasoning.
- Increased zero-shot QA accuracy from **50% to 77%** using targeted clarifications, validated via ROUGE-L.

Pratibha – Learning Management System - ReactJS, Firebase, TailwindCSS, ExpressJS

- Built and deployed a full-stack LMS platform used by **6000+ students and faculty** across 10+ departments, digitizing file submissions, grading, and feedback workflows reduced grading turnaround by **80%**.
- Implemented secure, role-based access and **SSO integration** with Google OAuth, optimized component rendering and API latency, reducing average page load time to under **300ms** via lazy loading and indexed Firestore queries.

Achievements and Certifications

- All India Rank 15 - IEEE Xtreme15.0 - 24 Hour programming contest.
- Google Developer Student Club - AI/ML Lead: Conducted Technical Bootcamps and Hackathons.
- **AWS Academy** - Cloud Foundations