Harshita Dooja Poojary

hpoojary@usc.edu | +1(213)-234-8695 | harshitapoojary | Publications

Analytical and passionate technologist with 4.5 years of experience in machine learning, model deployment, backend development, and CI/CD automation, with strong communication, organization, teamwork, and a proactive approach to learning and problem-solving.

EDUCATION

University of Southern California, Los Angeles ***May 2025***

Master's in Computer Science (Coursework-*Algorithms, Deep Learning, Artificial Intelligence, Machine Learning, Natural Language Processing*) **GPA: 3.81**

University of Mumbai, India ***May 2019***

Bachelor's in Computer Science Engineering (Coursework- *Software Engineering, Distributed Databases, Web Technologies*) **GPA: 3.83**

SKILLS

Core Programming Languages: Python

Other Programming Languages: JavaScript, C++, Java, HTML/CSS

Frameworks & Libraries: Node.js, AngularJs, PySpark, PyTorch, Keras, TensorFlow, OpenCV, Streamlit, Scikit-learn

Databases: MongoDB, Redis, MySQL, ElasticSearch, Milvus, Hadoop, DynamoDB

Tools & Platforms: Docker, Kubernetes, AWS, Azure, Jenkins, GitHub, GitLab, JIRA, Postman, Swagger, PowerBI, Tableau, Google Analytics, GCP

Functional Skills: Data Science, Software Engineering, Machine Learning, Deep Learning, Data Structures & Algorithms, Software Development, Large Language Models, Statistical Analysis, MLOps, Data Visualization, Data Analytics, Cross-Functional Collaboration, Artificial Intelligence

WORK EXPERIENCE

Data Science Intern ***Jun 2024 – Aug 2024***

Zero-True (Remote) New York, USA

- Conducted comprehensive data analysis on **10 distinct datasets** using **ML** and **NLP techniques**, employing the visual components of Zero-True to enhance client experience by enabling dynamic data filtering and visualization using Vue.js, Python (Scikit-learn, PyTorch)
- **Collaborated with the founders** to implement **5 additional UI components** from Vuetify, contributing to the development of a seamless Python and SQL integration for creating interactive reports and dashboards.
- **Optimized data processing for accelerated ML inference** within Zero-True notebooks, enabling seamless integration and efficient deployment of computationally intensive applications under restricted resources.

Student Researcher ***Jan 2024 – Aug 2024***

University of Southern California Los Angeles, USA

- Under **Prof. Seon Kim**, analyzed **object detection and classification models** (YOLOv5/YOLOv8) on 10k Streetview imagery from Google Images and in-house dataset using **data mining** for tent detection and counting, optimizing performance in **noisy urban conditions** using geospatial metadata.
- **Executed geospatial analysis at scale** by coding grid-based segmentation logic to quantify encampment density across **502.7 square miles** of Los Angeles, generating multi-intensity **heatmaps** using latitude-longitude clustering **across five intensity levels**.

Machine Learning Engineer ***Dec 2021 – Jul 2023***

Reliance Jio Mumbai, India

- **Engineered and optimized** RetinaFace detection models using **TensorRT FP16** precision on **T4 and A100 GPUs** using **C++ and Python**, reducing face registration time by **40%** and enabling real-time inference at scale across Jio's parallel distributed clusters.
- Automated end-to-end **CI/CD pipelines** for three ML systems—JioFace, Video Motion Detection, Number Plate Recognition—reducing model deployment time by **50%** and supporting rapid model iteration with automated MLOps.
- Implemented **mobile-based spoof detection** to prevent false facial verification using a photo or video with 90% accuracy by creating a spoof detection system with **Resnet Architecture in H2O.ai (no-code platform).**
- Led a team of interns to **curate spoof detection datasets**, conduct model evaluation experiments, and establish **reproducibility benchmarks** to validate model integrity and deployment readiness at scale.

Software Engineer ***Jun 2019 – Dec 2021***

Reliance Jio Mumbai, India

- **As a part of the world's largest mobile data company, collaborated** cross-functionally to build scalable microservices and REST APIs with SQL, NoSQL databases serving personalized ML-driven recommendations for 100M+ users, enhancing engagement by 20%.
- **Designed large-scale data ingestion pipelines**, integrating 250+ newspapers, 150+ live channels, and 800+ magazines across platforms, **scaling content delivery** in 13 languages and increasing the user base by **50%.**
- **Augmented system observability** and **SLA adherence** by implementing a full-stack monitoring suite using **Prometheus, Grafana, and Kibana**, introducing memory profiling and automated alerting that improved API availability and reliability by **20%.**

Software Engineer ***Jan 2019 – Jun 2019***

TaksyKraft (Remote) Hyderabad, India

- Built scalable **data pipelines** and **visualizations** for **10 large-scale clients** in collaboration with analysts and engineering teams.
- Designed and deployed **APIs** to deliver **real-time engagement predictions** and **ROI metrics**, enabling 10+ broadcasters and advertisers to optimize ad targeting.

PROJECTS

Unified Semantic Space for Multimodal Retrieval (PyTorch, Python, VectorDB)

- Analyzed multimodal LLMs, including **LLama-3,** to answer questions over frames for MIT videos and using **LangDB** Vector Databases to store frame embeddings and retrieve user queries.
- Implemented a user interface using **Streamlit** to integrate ML models, enabling users to upload videos and query the system interactively.

Analyzing NLP techniques on Low-Resource Languages(PyTorch)

- **Conducted idiom classification and metaphor classification** in the low-resource **Konkani** language using BERT-based transformers, investigating performance-preserving **attention head pruning.**

Comparison of Model Performance with Knowledge Distillation and Quantization (PyTorch, Python)

- **Fine-tuned LLama-3 model** with 4-bit quantization using **logit distillation** with LoRa (Low-Rank Adaptation).
- Computed benchmark against **GSM8k dataset with 8-shot learning** to analyze the performance against the original model with 8B parameters.