

# DARSHAN MAKWANA

[darshanmakwana412@gmail.com](mailto:darshanmakwana412@gmail.com) ◇ [GitHub](#) ◇ [LinkedIn](#) ◇ [Twitter](#)

WEBPAGE: <https://darshanmakwana412.github.io>

My name is Darshan Makwana and I am an AI engineer at [Riverline AI](#) building robust and scalable agentic eval systems. Prior to this I was an ML engineer at [Sprinklr](#) working on real time voice AI systems and improving their ASR inference. I was also a Research Assistant at [Aalto Vision Lab](#) working on [3D Gaussian Splatting](#) and also did a summer research internship at [TU Munich](#) on [quadruped robotics](#).

I graduated from [IIT Bombay](#) with a B.Tech in Mechanical Engineering and Dual Minor in Computer Science and Machine Learning. My interests span efficient ML serving systems, 3D reconstruction and endurance sports.



## EDUCATION

---

**Indian Institute of Technology, Bombay** 2021 - 2025  
B.Tech in Mechanical Engineering | Minor in Computer Science and Data Science  
Major CPI: 8.82/10 · Minor CPI: 10/10

- **Team Lead, Inter IIT Tech Meet 13.0** Dec 2024  
Led team of 13 for Adobe deepfake detection challenge to 2nd runner up
- **Teaching Assistant** Apr - Jun 2023  
PH 112 & MA 108 | Conducting weekly tutorials and grading assignments for 40+ students
- **Machine Learning Mentor** May - Jul 2023  
Mentoring 7 students during the summers to pick up pace in ML
- **Core Team Member, DAV** Jul 2022 - Apr 2023  
Data Analytics Team of the Undergraduate Academic Council
- **Web Coordinator** May 2022 - Jun 2023  
Developed websites for SARC events and alumni portals

**Aalto University, Finland** Jan - Jun 2025  
Semester Exchange, School of Science  
Ranked 3/380+ in Programming Parallel Computers Course  
[Letter of Recommendation](#)  
CPI: 10.0/10.0

## EXPERIENCE

---

**Riverline AI** Bengaluru, India  
*AI Engineer (Full-time)* May 2026 - Present  
Building end to end eval pipelines for voice agents, scoring live and recorded call sessions across compliance adherence, litigation constraints, and operational spec conformance. Developing automated agentic judges to flag violations and measure agent behavior across multi turn conversations.

## Sprinklr

Machine Learning Engineer (Full-time)

Gurugram, India  
Aug 2025 - Apr 2026

Built robust inference serving systems for realtime voice AI applications serving 10M+ monthly voice minutes. Architected changes around whisper to enable  $2.5\times$  increase in throughput with minimal hit on performance. Designed a scheduling strategy for ASR systems achieving 43% reduction in median E2E latency at high loads (Published at CAO @ ICLR 2026).

Won 2nd prize in company wide voice ai hackathon building an end to end voice agent prototype with real time multi-turn conversation handling and tool calling capabilities.

## Aalto Vision Lab

Research Assistant under *Matias Turkulainen* & *Juho Kannala*

Helsinki, Finland  
Jan - Jul 2025

Wrote multi GPU training scripts for feedforward gaussian splatting on SLURM clusters. Profiled and optimized training kernels for better GPU utilization. Benchmarked and ablated various optimization techniques for feedforward gaussian splatting.

## Sprinklr

Product Intern

Gurugram, India  
May - Jul 2024

Benchmarked audio codecs for use in STT and TTS applications. Built a prototype of a unified voice agent for customer support services. Received Pre Placement Offer (PPO).

## Technical University of Munich

Summer Research Intern | *Quadruped Robotics* under *Hongpeng Cao* & *Marco Cacca*

Munich, Germany  
Jun - Jul 2023

Developed system identification models for dynamics modelling of quadruped robots and implemented MPC based controllers and compared against model free deep RL approaches in simulation.

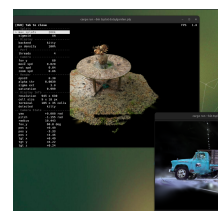
## PROJECTS

---

### Gaussian Splatting TUI | [tsplat](#)

Apr 2026

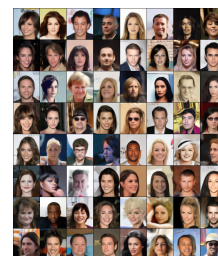
tsplat renders 3D Gaussian Splatting scenes directly in your terminal using Unicode half blocks or any other supported graphics protocol, works with xterm, kitty, ghostty. It's written in rust and is CPU only for now, doesn't require GPU or display servers and it even works over SSH.



### Gaussian Masked Autoencoders | [code](#)

Feb - Mar 2025

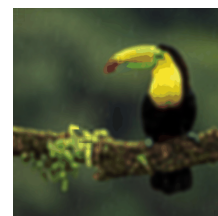
Open source implementation of the [Gaussian MAE](#) paper which leverages Gaussian splats as intermediate latent representation, enabling certain zero-shot capabilities like depth, splat grouping, etc.



### Minimal 2D Gaussian Splatting | [code](#)

Dec 2024

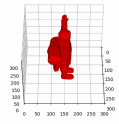
A minimal Implementation of gaussian splatting for 2D images in triton. Uses tile based differential rasterization similar to the original gaussian splatting paper but only works on 2D gaussian splats.



### Multi-View Reconstruction | [code](#) | [report](#)

Sep - Nov 2024

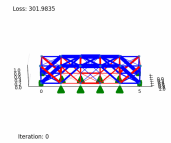
An implementation of volumetric graph cut algorithms on visual hulls for 3D reconstruction from multiple views. Additionally ray casting for shading is used to render the final scene. This was our final course project for ME 735.



### Structural Optimization | [code](#) | [report](#)

Jan - Apr 2024

A framework for structural optimization of trusses under loading. It uses a matrix formulation of structures and gradient based optimization to compute the optimal spatial configuration of the structure under any loading conditions.



### Fractal Curves for Tool Path Planning | [code](#) | [report](#)

Aug - Nov 2023

This is a PoC of using fractal curves for tool path planning in layered 3D printing. It implements a path planning algorithm that recursively decomposes a region into disjoint smaller fractals resulting in a continuous printing path for any layer.



## PUBLICATIONS

---

### Duration Aware Scheduling for ASR Serving Under Workload Drift

[CAO @ ICLR 2026](#)

Darshan Makwana, Yash Jogi, Harsh Kotta, Aayush Kubba.

### COVID Self diagnosis classification using BERT and LightGBM models

[SMM4H 2023](#)

R Chavda, Darshan Makwana, et al.