www.anujasaini.com
anujasaini96@gmail.com
+91 75788 25002
github.com/anujasaini
linkedin.com/in/anujasaini

Anuja Saini

Applied Research Engineer · Computer Vision & Applied Machine Learning

Applied research engineer with seven years of taking machine learning from first experiment to production across 3D perception, computer vision, forecasting and causal inference. Built a company's first LiDAR product from zero to paying customers on three continents in 14 months, publishing three papers along the way; earlier, built forecasting systems at Salesforce that serve 1B+ predictions a month and retained $15M+ in revenue. Looking for an early-stage team where research has to ship.

EXPERIENCE

AiDash · Senior Data Scientist, 3D Perception Research — Dec 2024 – Present

- Owned LiDAR end to end, from research through platform to customer delivery: first experiment to deployed pipeline in 14 months, now serving 500,000+ line miles of utility network across the US, Canada and Europe.
- **TripleInjection** (ICLR 2027): a small trainable network grafted onto a frozen PTv3 backbone, trained with a rare-class recipe and multi-density augmentation (full density → 5 pts/m^2). Wire IoU 0.96 on DALES, level with the strongest published baseline, and the first reported pole/wire results below 12 pts/m^2. On real utility scans it lifts wire IoU from 0.10 to 0.84 and transfers across sensors without retraining, replacing paid survey flights with free public LiDAR.
- **Grid topology** (BMVC 2026 · IGARSS 2026 oral): a dual-head ResUNet-FPN on DSM rasters turns segmented points into a connected pole/span graph with each conductor modelled individually; locates 85% of poles against the utility's GIS records on low-density USGS LiDAR, replacing manual field verification.
- **Production delivery**: designed and deployed a five-stage inference pipeline on AWS SageMaker (segmentation → topology heatmap → polygonisation → metadata tagging) as independent scale-to-zero endpoints. Terraform, commit-pinned containers and MLflow model aliases mean a new model ships as a config change; run at 50,000+ line miles in a single job.
- **Evaluation and data**: built the cross-dataset benchmark that decides what ships (four public LiDAR benchmarks plus six utility datasets totalling 80M points, with a sparse-3D annotation strategy for ground truth) and the customer-facing accuracy report against utility GIS records at graded match tolerances, now the acceptance criterion for new deployments.
- **Imagery-only fallback**: adapted Depth Anything V2 to canopy-height data for terrain and vegetation reconstruction where no LiDAR exists, extending coverage to imagery-only regions.

Salesforce · Senior Data Scientist — May 2021 – Nov 2024

- Built churn prediction for the full customer base: LSTM models serving 1B+ forecasts a month, with SHAP and causal inference to isolate the drivers of attrition so account teams act on causes, not correlations.
- Productionised retraining and scoring end to end in Airflow, cutting model refresh from days to hours and removing the manual monthly run.
- Designed the customer health scorecard (engagement, adoption and retention per account) with a recommender that prescribes the training most likely to lift each customer's productivity; adopted by success teams as the shared view of every account: $15M+ in revenue retained, 97K+ customer hours saved.

EXL Service · Consultant — Sep 2019 – May 2021

- Started the practice's computer-vision work from scratch: building-damage detection and severity grading from pre- and post-disaster satellite imagery, built end to end from scene tiling and labelling to a change-detection CNN and severity classifier.
- Designed the measurement framework for a retail-tech client's direct-mail campaigns: weekly KPIs (activation, response rate, incremental spend) with statistical tests, used to decide campaign spend.

PUBLICATIONS

ICLR 2027 — Sparse Wires, Robust Models: Density-Robust Segmentation of Poles and Wires from Open Aerial LiDAR. Second author · Under review

BMVC 2026 — Robust Utility Networks Registration with Powerline Segmentation on Noisy Point Clouds. A. Saini, R. Jain, V. Jain

IGARSS 2026 — Learning-Based 3D Reconstruction of Power Networks from Aerial Point Clouds. R. Jain, A. Saini, V. Jain · Oral

TECHNICAL

Perception — 2D/3D semantic segmentation · detection · change detection · monocular depth · vision foundation models (fine-tuning, zero-shot) · ViTs, CNNs · aerial & satellite imagery

3D & LiDAR — Point Transformer v3 · KPConv · PointNet++ · point segmentation · Gaussian splatting · rasterisation

ML — Rare-class & imbalanced learning · time-series forecasting · causal inference · explainability (SHAP) · cross-dataset benchmarking

Engineering — Python · SQL · PyTorch · TensorFlow · OpenCV · XGBoost / LightGBM · Spark · AWS (SageMaker, S3, EC2) · Airflow · Docker · Terraform · MLflow · CI/CD

EDUCATION

IIT Guwahati · B.Tech, Chemical Engineering, 2019

Stanford University · AI Professional Program: NLP with Deep Learning, NLU, ML with Graphs

Future University Hakodate, Japan · Research Intern, 2018: CNN hand detection in egocentric video

GrowthX · Growth capstone, top 4 of cohort: plan to double Yellow.ai's India revenue, $40M → $80M