

Tharun Krishna

tharunkrishna1611@gmail.com | github.com/Tharun-bot | tharun-port.vercel.app | linkedin.com/in/tharun-krishna

EDUCATION

Indian Institute of Technology Madras

BS in Data Science and Applications

Chennai, India

07/2023 – 04/2027

National Institute of Technology Karnataka

B.Tech in Mining Engineering

Surathkal, India

07/2023 – 04/2027

EXPERIENCE

ML Research Intern

05/2026 – 07/2026

National Institute of Technology Karnataka

Surathkal, India

- Analyzed field blast-monitoring data and engineered physics-informed features to model **blast-induced ground vibration (PPV)**, running **exploratory data analysis**, correlation studies, and hypothesis tests to characterize the underlying attenuation relationships
- Evaluated multiple regression models using **learning curves** and cross-validation to diagnose bias/variance trade-offs, selecting a model resilient to overfitting on a limited field dataset

AI Research Intern

01/2025 – 03/2025

Scalar Labs

Surathkal, India

- Designed a **real-time time series forecasting system** that converts streaming numerical data into **natural language embeddings** and predicts future values using **pretrained, frozen LLMs (PyTorch, Transformers)**
- Implemented **embedding alignment** techniques to bridge numerical and language-based representations, enabling efficient **next-K-step forecasting** without fine-tuning the base LLM

Machine Learning Engineer Intern

05/2024 – 07/2024

Vaidsys Technologies

Remote, India

- Developed a **real-time anomaly detection pipeline** for streaming API and IoT sensor data, **cutting the false-positive alert rate by 35%** through feature engineering and adaptive threshold calibration
- Deployed **ML inference services on AWS EC2** with **Docker**, achieving **sub-100ms prediction latency** with **S3-backed model artifact versioning**

PROJECTS

BlastPred | Python, XGBoost, scikit-learn, SHAP, Optuna, SciPy | github.com/Tharun-bot/blast-induced-ppv-pred

- Built a **physics-informed ML pipeline** on **200 field blast-monitoring records** that predicts ground vibration along **three orthogonal directions (V/L/T)** instead of a single resultant PPV, engineering features from the Holmberg–Persson attenuation relation (scaled distance, log-transforms, charge-distance interactions)
- Tuned a **stacking ensemble (XGBoost + SVR, Ridge meta-learner)** with **Optuna** Bayesian search, improving **R^2 from 0.750 to 0.796–0.819** and cutting **RMSE by 16–20%** versus the empirical baseline; used **SHAP** and Spearman hypothesis testing to confirm statistically significant directional anisotropy ($p=0.004$), improving reconstructed-PPV accuracy by **4.14%** over direct prediction

ClinicalLens | Python, Lifelines, scikit-learn, Streamlit | github.com/Tharun-bot/clinicalLens

- Built a **survival analysis and causal inference pipeline** on ICU clinical data using **Kaplan-Meier estimation**, **Cox Proportional Hazards** models, and **Inverse Probability of Treatment Weighting (IPTW)** to correct for confounding bias
- Identified and corrected a **Simpson's Paradox** in ventilation outcomes: naive analysis showed **HR $\approx$ 1.35** (appearing harmful), while **IPTW-adjustment recovered the true protective effect (HR $\approx$ 0.75)**, delivered via an interactive **Streamlit** dashboard

TECHNICAL SKILLS

Languages: Python, SQL, R, Go, JavaScript

Data Science: pandas, NumPy, scikit-learn, XGBoost, SHAP, Optuna, Lifelines, SciPy

Visualization & Tools: Matplotlib, Streamlit, Jupyter, Git

Concepts: Regression & Ensemble Modeling, Survival Analysis, Causal Inference, Time Series Forecasting

Other: Docker, REST APIs, Linux

CERTIFICATIONS AND EXTRACURRICULAR

Solved **400+** problems on LeetCode

Oracle Cloud Infrastructure 2025 Certified Generative AI Professional | Certificate

Oracle Cloud Infrastructure 2025 Certified AI Foundations Associate | Certificate