Jithendra Puppala

+1 (424) 944-2147 | jithendra.mail.me@gmail.com | New York, NY
Jithendra.com | github.com/jithendra1798 | linkedin.com/in/jithendra-siddartha

Summary

Data Scientist bridging industry and research: 2 years deploying ML at scale (150K+ devices, $10M+ impact) plus hands-on experience with Vision Transformers, scaling laws, and GPU-optimized training. MS CS at NYU. Seeking Summer 2026 AI/ML/Data Science internship.

Education

New York University — New York, NY
Master of Science in Computer Science | GPA: 3.67 — Expected: 05/2027
- Coursework: Machine Learning, Computer Vision, Big Data

National Institute of Technology Karnataka — Surathkal, India
Bachelor of Technology in Computer Science and Engineering — 05/2023
- Coursework: Data Structures and Algorithms, Digital Image Processing, Probability & Statistics

Skills

- **Languages & Core ML**: Python, SQL, PyTorch, TensorFlow, Scikit-learn, XGBoost, Pandas, NumPy
- **Deep Learning**: Vision Transformers (ViT, Swin, ConvNeXt), CNNs, LLMs/Transformers, CUDA, Mixed Precision
- **Computer Vision**: OpenCV, YOLO, Object Detection & Tracking, Image Classification, Multi-View Fusion
- **MLOps & Cloud**: Docker, FastAPI, REST APIs, AWS (EC2, S3), GCP, Azure, CI/CD, Git, Linux
- **Analytics**: A/B Testing, Statistical Analysis, Data Wrangling, Model Evaluation & Optimization

Work Experience

Jio Platforms Limited — Bengaluru, India
Data Scientist — 06/2023 – 07/2025
- Developed production ML classification system achieving **91% precision** and **95% recall** (up from 68% and 76%), serving real-time predictions to **150K+ IoT sensors** via FastAPI microservices
- Engineered ARIMA-based anomaly detection enabling **4-hour earlier predictions**, improving success rates by **28%** and driving **$10M+** annual savings
- Built signal processing pipeline using Butterworth filters, improving activity classification accuracy from 87% to 94% and reducing false positive alerts by 28%
- Optimized end-to-end ML retraining infrastructure: reduced model retraining time from **18 hours to 2 hours** (89% reduction) and data processing from 3 hours to 7 minutes (96% reduction)

Projects

GeoGuessr US State & GPS Prediction — 11/2025 – 01/2026
- **Ranked 1st among 50+ teams on Kaggle** by developing **multi-view attention architecture** fusing 4 street-view images for 50-state classification + GPS regression, fine-tuning 7 Vision Transformer backbones (ViT-CLIP, EVA02, Swin, BEiT, ConvNeXt)
- Achieved **96%+ competition score** via **weighted ensemble** with ablation-informed model selection and **Test-Time Augmentation (TTA)**, improving over best single-model baseline by 3%
- Engineered GPU-optimized data pipeline with **tensor caching** and **on-GPU augmentation** (MixUp, CutMix), reducing epoch time by **40%** on A100 80GB

Scaling Laws for Language Models on Symbolic Music Data — 11/2025 – 01/2026
- Investigated **neural scaling laws** on symbolic music by training **9 decoder-only Transformer and LSTM models** (844K–201M parameters) on **100M+ tokens** of ABC notation from 54K+ folk tunes
- Demonstrated **1.2× better scaling behavior in Transformers than LSTMs**, with a **60%** reduction in validation loss when scaling from 844K to 201M parameters
- Optimized training on **A100-80GB GPU** using **mixed-precision (bfloat16)**, Flash Attention, achieving **1.28 test perplexity** and generating syntactically valid playable music

Multi-Object Tracking System — 08/2025 – 10/2025
- Built YOLOv10 + DeepSORT tracking webapp with async CUDA batching, reducing inference latency by 50%
- Deployed on Azure App Service (Docker + CI/CD) with 90% storage reduction via TTL cleanup and streaming I/O

Leadership & Achievements

- **Harvard CELP Finalist (Top 114 of 10 000+)** - selected for Harvard's global emerging leadership program
- JEE Advanced Rank: **6111** - Top **0.5% of 1.15M** candidates in India's most competitive engineering exam