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

MS CS at NYU with 2+ years as a Data Scientist at Jio Platforms, building production ML systems serving 150K+ IoT devices ($10M+ impact). 1st place in Kaggle computer vision (50+ teams), 3rd place in tabular ML (100+ teams) competitions. Currently leading AI-powered RAG and workflow automation at NYU.

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
- **Work Authorization:** US work-eligible via CPT (internship) and STEM OPT (3 years post-graduation)

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, Database Systems

Skills

- **Languages & Core ML**: Python, SQL, PyTorch, TensorFlow, Scikit-learn, XGBoost, LightGBM, Hugging Face, LangChain
- **Deep Learning**: Vision Transformers (ViT, Swin, ConvNeXt), CNNs, LLMs/Transformers, CUDA, Mixed Precision, Flash Attention, Natural Language Processing
- **Computer Vision**: OpenCV, YOLO, Object Detection & Tracking, Image Classification, Multi-View Fusion
- **MLOps & Cloud**: Docker, FastAPI, REST APIs, AWS (EC2, S3), GCP, Azure, PySpark, MLflow, CI/CD, Git, Linux
- **Analytics**: A/B Testing, Hypothesis Testing, Feature Engineering, Causal Inference, Data Wrangling, Pandas, NumPy

Work Experience

New York University – Office of Faculty Affairs — New York, NY
Lead Developer – Data Science & AI — 02/2026 – Present
- Built **RAG**-based intelligent ticketing prototype with LangChain, ChromaDB, and sentence-transformer embeddings for faculty information document querying with retrieval accuracy of **94%**, scaling to **automate resolution of 900+ annual workflow requests** across **15+ administrative processes**
- Developing automated **workflow translation layer** using Python, converting legacy manual processes into structured digital pipelines with **real-time status dashboards** for operational monitoring for **300+ faculty**

Jio Platforms Limited — Bengaluru, India
Data Scientist — 06/2023 – 07/2025
- Owned end-to-end ML classification pipeline serving real-time predictions to **150K+ IoT devices** via FastAPI microservices, improving **precision from 68% to 91%** and **recall from 76% to 95%**
- Built ARIMA-based anomaly detection that **reduced alert latency from 12 hours to 4 hours** (67% faster), improving field response success rates by **28%** and driving **$10M+** in annual savings
- Re-architected ML retraining infrastructure: model retraining from **18 hours to 2 hours** (89% reduction), data processing (ETL) from **3 hours to 7 minutes** (96% reduction)

Projects

Multi-View Vision Transformers for Geospatial Prediction *(1st Place, Kaggle - 50+ Teams)* — 11/2025 – 01/2026
- Developed multi-view attention architecture fusing 4 street-view images for 50-state classification + GPS regression, fine-tuning 7 ViT backbones (ViT-CLIP, EVA02, Swin, BEiT, ConvNeXt)
- Achieved **96%+ score** via weighted ensemble with ablation-informed model selection and Test-Time Augmentation (TTA)
- Engineered GPU-optimized pipeline with **tensor caching** and **on-GPU** augmentation (MixUp, CutMix), reducing epoch time by **40%** on A100

Neural Scaling Laws: Training Transformers up to 201M Parameters — 10/2025 – 12/2025
- Trained **9 decoder-only Transformer and LSTM** models (**844K - 201M params**) on **100M+ tokens** of structured symbolic sequences, demonstrating **1.2× better scaling** in Transformers with **60%** lower validation loss
- Optimized training on **A100-80GB** using mixed-precision (bfloat16) and Flash Attention, achieving **1.28 test perplexity**

Pairwise Compatibility Prediction from Professional Profiles *(3rd Place, Kaggle - 100+ Teams)* — 02/2026
- **Ranked 3rd among 100+ teams** predicting pairwise compatibility scores from **mixed text and categorical profiles** in a 3-day ML hackathon (Codefest'26, IIT BHU Varanasi)
- Reverse-engineered scoring formula through systematic combinatorial feature analysis across **8 attribute fields**, achieving 100% prediction accuracy using **optimized Jaccard similarity** over parsed text and categorical features

Leadership & Achievements

- **CV Engineer, NYU Robomasters** - Developing perception and object detection systems for competitive robotics
- **Harvard CELP Finalist (Top 114 of 10 000+)** | **JEE Advanced Rank: 6111** - Top **0.5%** of 1.15M candidates
- **Community:** Campus Director, Millennium Fellowship (led 30 fellows) | 3,300+ LinkedIn followers