Domenico Ceresa

San Francisco Bay Area | (925) 407-7855 | ceresanico@gmail.com | nico-ceresa.com | linkedin.com/in/domenicoceresa

OVERVIEW

Data Scientist with a strong foundation in machine learning and AI, backed by an M.S. in Computer Science from Georgia Tech and 3+ years of applied industry experience. Skilled in building ML pipelines, solving business problems, and building automations.

EDUCATION

Georgia Institute of Technology | Atlanta, GA — Aug 2025 - Jun 2027

M.S., Computer Science

- Specialization: Artificial Intelligence
- Courses: Robotics & AI Technology, Machine Learning

University of Oregon | Eugene, OR — Oct 2019 - Jun 2023

B.S., Data Science & B.S., Economics — *Major GPA: 3.88/4.00*

- Minors: Computer Science, Mathematics
- Relevant Courses: Statistics for Data Science, Machine Learning, Linear Algebra, Data Ethics

WORK EXPERIENCE

BPM LLP | San Francisco, CA

Data Analyst — Jun 2024 - Apr 2026

- Engineered an NLP pipeline integrating Azure Cognitive Services and GPT-4 to summarize 100+ page transfer pricing reports, reducing analyst preparation time by ~98% (4 hours to ~5 minutes)
- Designed and delivered Power BI dashboards tracking operational metrics giving leadership a view of firm performance to drive resourcing and financial decisions
- Applied fuzzy matching algorithms to reconcile invoice line items against project estimates, generating automated PDF outputs and reducing processing time from ~5 hours to under 1 minute
- Built a simulation model to optimize intercompany renewable energy transfer pricing to maximize profit through Canadian carbon credit monetization

Data Analyst Intern — Apr 2023 - Jun 2024

- Developed Power BI dashboards to surface key metrics and employee performance, providing clientele visibility into their data
- Conducted research and synthesized findings to support senior analysts on client engagements, contributing to data-driven recommendations and deliverables
- Produced internal and external marketing materials to increase awareness of the data team's capabilities, driving cross-functional engagement and stakeholder adoption
- Collaborated cross-functionally with leadership and business teams to communicate data insights and support strategic decision-making

Univ. of Oregon, Data Science Dept. | Eugene, OR

Data Science Learning Assistant — Dec 2022 - March 2023

- Facilitated weekly office hours supporting 50+ students
- Supported lessons, testing, and grading of student work
- Helped students gain a better understanding of regression, KNN, gradient descent, decision trees, bootstrapping, and more

PROJECTS

CruxCam | Computer Vision

- Built a real-time pose estimation app using MediaPipe and OpenCV to analyze climbing video and score technique efficiency
- Engineered a joint angle and center-of-mass calculation pipeline to generate an efficiency score based on body position
- Deployed as an interactive Streamlit web app with video output, live efficiency overlays, and adjustable detection thresholds
- Architected clean separation between pose analysis and video I/O logic via a modular core package

Super Resolution Paper Implementation | Computer Vision

- Implemented a PyTorch-based FSRCNN super-resolution model and trained it end-to-end on custom image datasets
- Built reusable data loading and preprocessing pipelines to support efficient training and evaluation
- Evaluated model performance using PSNR and achieved results consistent with published benchmarks
- Deployed the model via a Flask service and containerized the system with Docker for reproducibility

SpaceX CRS-10 Telemetry Analysis | Time Series/Anomaly Detection

- Investigated three unsupervised anomaly detection methods (Z-score, IQR, Rolling Z-score) on high-frequency SpaceX CRS-10 rocket telemetry, using 5 known flight events as ground truth
- Built a cluster-based evaluation framework to group detected anomalies and score each method on precision, recall, and F1, enabling direct comparison across techniques
- Discovered through grid search that the optimal Z-score threshold (0.7) was far below the conventional 2.0–2.5 range, revealing the signal's compressed variance as a key driver
- Concluded that locally-adaptive methods consistently outperform global statistics on non-stationary signals, a finding with direct implications for real-world sensor data pipelines

SKILLS

- *Languages & Data:* Python, SQL, Pandas, NumPy
- *ML / AI:* Scikit-learn, PyTorch, TensorFlow
- *Computer Vision:* OpenCV, MediaPipe
- *Tools & Visualization:* Git, Docker, Flask, Power BI, Matplotlib, Seaborn