

Diego Arteaga

(661) 219-4382 | darteaga520@gmail.com | darteaga.dev | linkedin.com/in/arteagadiego

Education

| | |
|---|--|
| California State University, Northridge <i>Master of Science in Data Science</i> <i>Bachelor of Science in Computer Science (Summa Cum Laude)</i> | Northridge, CA May 2027 May 2025 |
| <ul style="list-style-type: none">Relevant Coursework in: Data Structures, Algorithm Design & Analysis, Software Engineering, Machine Learning, Deep Learning, Data Mining, Database System Design, Python Programming for Data Analytics, Probability & Stats | |

Experience

| | |
|--|--|
| Unclassified Paid Intern <i>Los Angeles Valley College</i> | Sept 2025 - Present Los Angeles, CA |
| <ul style="list-style-type: none">Maintained and updated confidential student records from PeopleSoft in Microsoft Excel, supporting outreach and recruitment reporting needs while ensuring data integrityRepresented the college at high schools and community youth centers, assisting prospective students | |
| Social Media Manager & Video Editor <i>Self-Employed</i> | Nov 2021 - Present Los Angeles, CA |
| <ul style="list-style-type: none">Analyzed YouTube Studio & TikTok analytics to design and execute A/B tests on content strategy, increasing revenue by 22% and audience engagement by 40%Produced and edited engaging content for YouTube, TikTok, and Instagram, optimizing workflow efficiency using Ableton Live, Adobe Creative Cloud tools (Premiere Pro, After Effects), and Photopea, improving turnaround timeBuilt Google Sheets dashboards to track KPIs and make data-driven content strategy decisions | |
| Undergraduate Research Intern <i>California State University, Northridge</i> | June 2025 - Aug 2025 Northridge, CA |
| <ul style="list-style-type: none">Implemented an anomaly detection pipeline on IoT healthcare datasets using Python and PyTorchPreprocessed data, optimized models, and collaborated in a team of 5, delivering reproducible results, technical documentation, and a presentation for stakeholders | |

Projects

| | |
|---|--|
| Japanese-to-English Translation Model <i>Python, PyTorch, Moses, MeCab, SentencePiece</i> | <ul style="list-style-type: none">Developed a neural machine translation system using a 3,000,000 sentence parallel corpus to improve Japanese-to-English translation accuracy in scientific and technical domains, such as biology and medicinePerformed multilingual preprocessing with Moses, MeCab, and joint SentencePiece subword tokenization |
| Hybrid IoMT Anomaly Detection System <i>Python, PyTorch, Scikit-learn, Google Colab</i> | <ul style="list-style-type: none">Preprocessed and cleaned 4 public healthcare IoT datasets into a reproducible analytical pipelineEngineered temporal and graph-based features to detect abnormal network behaviorProject recognition: Best Paper Presentation, IEEE IEMCON 2025 |
| Yoda NLP Model <i>Python, Hugging Face, PyTorch</i> | <ul style="list-style-type: none">Fine-tuned a pretrained natural language processing model using LoRA on a curated Star Wars dialogue corpus to mimic Yoda's non-standard syntactic structures.Evaluated adaptation using perplexity and TF-IDF cosine similarity against held-out dialogue. |
| Agricultural IoT & Machine Learning Platform <i>React, Bootstrap, Node.js, Python, Postman, Firebase</i> | <ul style="list-style-type: none">Designed a real-time dashboard using React and Bootstrap to visualize environmental sensor data from Firebase within a team of 12 students via GitHub and Jira, integrating and testing Python backend APIsCollaborated on machine learning models to correlate environmental data with plant sickness and crop growth1st Place, CSUNposium 2025 - top presentation selected across multiple majors and departments |

Skills

| |
|---|
| Programming Languages: Python, Java, JavaScript, SQL (MySQL, Google BigQuery), HTML/CSS |
| Frameworks/Libraries: React, Node.js, Express.js, NumPy, Pandas, Matplotlib, Scikit-learn, PyTorch, TensorFlow |
| Developer Tools: Git, GitHub, VS Code, Jira, Jupyter Notebook, Google Colab, NPM, AWS, Tableau |
| Certifications: Google Data Analytics Professional Certificate (Coursera, 2025) |
| Languages: English, Spanish, Japanese |