

# Diego Arteaga

📞 (661) 219-4382 | ✉️ [darteaga520@gmail.com](mailto:darteaga520@gmail.com) | 🌐 [darteaga.dev](http://darteaga.dev) | 🔗 [linkedin.com/in/arteagadieago](https://linkedin.com/in/arteagadieago)

## Education

### California State University, Northridge

Northridge, CA

Master of Science in Data Science

May 2027

Bachelor of Science in Computer Science (Summa Cum Laude)

May 2025

- **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

Sept 2025 - Present

Los Angeles Valley College

Los Angeles, CA

- Maintained and updated confidential student records from PeopleSoft in **Microsoft Excel**, supporting outreach and recruitment reporting needs while ensuring data integrity
- Represented the college at high schools and community youth centers, assisting prospective students

### Social Media Manager & Video Editor

Nov 2021 - Present

Self-Employed

Los Angeles, CA

- 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 time
- Built **Google Sheets** dashboards to track KPIs and make data-driven content strategy decisions

### Undergraduate Research Intern

June 2025 - Aug 2025

California State University, Northridge

Northridge, CA

- Implemented an anomaly detection pipeline on IoT healthcare datasets using Python and PyTorch
- Preprocessed 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 | Python, PyTorch, Moses, MeCab, SentencePiece

- 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 medicine
- Performed multilingual preprocessing with Moses, MeCab, and joint SentencePiece subword tokenization

### Hybrid IoMT Anomaly Detection System | Python, PyTorch, Scikit-learn, Google Colab

- Preprocessed and cleaned 4 public healthcare IoT datasets into a reproducible analytical pipeline
- Engineered temporal and graph-based features to detect abnormal network behavior
- Project recognition: **Best Paper Presentation, IEEE IEMCON 2025**

### Yoda NLP Model | Python, Hugging Face, PyTorch

- 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 | React, Bootstrap, Node.js, Python, Postman, Firebase

- 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 APIs**
- Collaborated on machine learning models to correlate environmental data with plant sickness and crop growth
- **1st 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