Jeremy Cheng

781-454-6369 | jec061@ucsd.edu | https://www.linkedin.com/in/jeremy-cheng | https://github.com/jeremyjcheng

Technical Skills

Languages: Python, Java, SQL, JavaScript, HTML/CSS
Skills: Data Analysis, Machine Learning, Natural Language Processing, Time Series Forecasting, API Integration
Tools: Git, Flask, PostgreSQL, Pandas, PyTorch, Scikit-learn, NumPy
Concepts: Data Engineering, Model Deployment, Feature Engineering, Fullstack Development, Recommendation Systems

Education

University of California, San Diego — La Jolla, CA
Bachelor of Science in Data Science — Expected June 2027

Work Experience

Software Engineering Intern — June 2025 - September 2025
Specter Aerospace — Peabody, MA

- Coordinate with test operators to develop **new features** for existing **facility control GUIs**
- Aid in the design and usability of **data dashboards** and **facility HMIs**
- Work with **propulsion** and **test engineers** in choosing charts and figures to generate and display
- Manage **database systems** for the **PLC** to **data server pipelines**

Data Science Intern — July 2024 - September 2024
Fizz Social — Palo Alto, CA

- Developed a news indexing system using **keyword extraction** and **sentence embeddings** to retrieve and rank relevant articles from NewsAPI and Google News.
- Refined search queries by implementing **query expansion** and **similarity search** mechanisms, improving the precision and relevance of News retrieval based on user input.
- Analyzed and visualized data, providing insights into user engagement and content relevancy through metrics including **NDCG (84%)**, **MAP (73%)**, and **MRR (83%)**.

Projects

EcoNaut — April 2025 – Present
Python, XGBoost, TimeGPT, Prophet, Pandas, PostgreSQL

- Designed and deployed a **full-stack anomaly detection pipeline** integrating ARGO float sensors and satellite data to monitor marine ecosystems.
- Implemented **multivariate time series forecasting** with **Facebook Prophet** and **TimeGPT**, enabling early detection of high-risk bloom conditions.
- Engineered oceanographic features to build an **XGBoost classifier** with **87% accuracy** in identifying algal bloom events.
- Structured a **PostgreSQL database** to support scalable querying, real-time model output logging, and future dashboard integration.

Sea Anomalies Detector — January 2025
PyTorch, Scikit-learn, Pandas

- Built a **fully connected neural network** in PyTorch during a hackathon to identify anomalies in marine float data.
- Designed the model architecture with **ReLU activations** and **30% dropout** to prevent overfitting and improve generalization.
- Performed dimensionality reduction and iterative tuning to enhance anomaly classification performance on time series inputs.

Pirate Predictor — December 2024
PyTorch, Flask

- Built an end-to-end **image classification system** to identify One Piece characters using a PyTorch-based deep learning model.
- Applied **transfer learning** by fine-tuning a pre-trained convolutional neural network on a custom-labeled character dataset.
- Developed a **Flask web app** with a lightweight interface to allow users to upload images and receive real-time predictions.

Beach Recommender — September 2024
Flask, Scikit-learn, Pandas

- Developed a Flask-based web application to recommend nearby beaches based on user location and surfing ability, integrating Google Maps API for geolocation and Storm Glass API for marine weather data retrieval.
- Implemented **ARIMA time series forecasting** to predict surf conditions, including wave height, wind speed, and swell, providing ranked beach recommendations tailored to user preferences.
- Engineered a feature pipeline combining real-time environmental inputs with user-level metadata to support adaptive recommendations.