ZHENGHAO GONG

george.zhenghaogong@gmail.com | (+1) 858-289-8748 | https://github.com/georgong

EDUCATION

University of California San Diego

Bachelor of Science in Data Science June 2026

- **Major GPA 3.9**
- **CourseWork:** Data Structure, Prob & Stat, Linear Algebra, Calculus, Machine Learning, DataFrame, DS principle, Data Visualization, Representation Learning, Probability Theory, Statistics, Deep Learning

SKILLS

Programming Language: python, c, javascript, html, css, rust
Data Science Framework: pandas & polars, scikit-learn, xgboost, lightgbm, catboost
Data Visualization: matplotlib, seaborn, plotly, d3.js
Deep Learning: pytorch, tensorflow, RNN, CNN, GAN, VAE, transformers.
Frontend-backend Interaction: flask, sanic
Math Base: Linear Algebra, Calculus, Probability theory, Convex Optimization, hypothesis test
Principle of Machine Learning Algorithm: Tree model, Linear model, Bayes net, Clustering, Decomposition.

EXPERIENCE & PROJECT

Image Inpainting Research: Auxiliary Loss Adaption for Image Inpainting
Group Member | Deep Learning, Computer Vision

- Contributed to the development of neural network architecture.
- Conducted research on the principle and innovation of auxiliary loss adaptation for Image Inpainting.
- Participated in the process of hyperparameter tuning, training, and comparative experiments among models.
- Contributed to the completion of the paper by creating and adding illustrations of the model architecture and algorithm principles.

Kaggle - Predict Energy Behavior of Prosumers
Individual | Machine Learning, Data Cleaning, Data Analysis, Feature Engineering, Time Series Data

- Conducted comprehensive data cleaning and Exploratory Data Analysis (EDA) utilizing polars.
- Executed feature engineering on Time-Series Data including STL decomposition
- Utilized LightGBM, XGBoost, and model stacking algorithms to forecast electricity consumption and production.

Github Project: The frontend visualization of model tuning
Project Leader | performance evaluation, hyperparameter tuning, frontend-backend Interaction

- Visualized the impact of hyperparameters on model performance and evaluated feature selections using appropriate metrics.
- Established a local server using Flask to facilitate frontend-backend interaction and implemented asynchronous updates of web pages via jQuery, HTML components, and JavaScript code.

Course Project: LoL esports match outcome prediction
Project Leader | Explory Data Analysis, Machine Learning, Decomposition, feature engineering

- Conducted Exploratory Data Analysis (EDA) to identify pertinent features within datasets.
- Applied PCA and Bayesian target encoding to address the curse of dimensionality and develop models from high cardinality categorical features.

LLM Application Internship
Group Member | Large Language Model, Prompting Engineering, Backend Server, Retrieval-Augmented Generation

- Developed large-scale model applications for deep learning video understanding and Retrieval-Augmented Generation (RAG) systems.
- Led the design and construction of a pipeline for the RAG system, optimizing response speed and inference performance using techniques like hybrid retrieval.
- Built a backend server and designed efficient parallel APIs to enhance system scalability and performance.
- Integrated a voice Q&A feature powered by local deep learning models, expanding the application's functionality.

LLM based web agent
Project Leader | Large Language Model, prompting engineering, computer vision

- Skilled in using Selenium to implement browser proxy automation tasks, including dynamic page interaction and data extraction.
- Proficient in using OpenCV for background subtraction and boundary detection, enabling efficient target detection and feature extraction.
- Integrated large language models (LLMs) into the proxy system to automate complex tasks and optimize inference speed and accuracy.