# Youcheng Taing

951-992-0563 | youchengtaing9999@gmail.com | github.com/Youcheng9 | linkedin.com/in/youcheng-taing-03423226a

## EDUCATION

**California State Polytechnic University**  **Pomona, CA**
*Bachelor of Science, Computer Science*  **Fall 2026**
- GPA: 3.90 / 4.0
- Relevant/Upcoming Courses: Statistics & Probability, Machine Learning, Data Visualization, Numerical Methods

## SKILLS & CERTIFICATES

**Programming & Tools**: Python, SQL, C++, Git/GitHub, Jupyter Notebook, Streamlit
**Data**: Pandas, NumPy, Matplotlib, Seaborn, Information Visualization
**ML Domains**: Scikit-learn, PyTorch, TensorFlow, NLP, Computer Vision (YOLOv8)
**Certificates: Cloud Computing 101**

## EXPERIENCE

**CPP AIoT Laboratory, Qcar Research Assistant | Python**  **January 2026 - Present**
- Collect, clean, and label 1,000+ driving images for supervised learning and model validation
- Develop and evaluate computer vision pipelines for object and lane detection on the Quanser QCar platform, achieving approximately 85% detection accuracy
- Measure model performance using precision, recall, and inference speed, achieving real-time inference

**CPP Autonomous Vehicle Laboratory, Member | PyTorch, YOLOv8**  **August 2025 - Present**
- Trained and evaluated YOLOv8 models on 10,000+ thermal datasets for object detection tasks, improving detection accuracy by 20%
- Tuned model hyperparameters including learning rate, batch size, and early stopping to reduce overfitting and stabilize training
- Designed data preprocessing and augmentation pipelines, increasing precision and recall by 10%
- Analyzed model performance using different metrics including training/validation loss, mAP50, mAP50-95, precision, recall, and comparative visualizations to support experimental conclusions

**Chaffey College Success Center, Tutor | Python, C++**  **October 2023 - December 2024**
- Tutored students in computer science and computer information systems courses using in-depth reviews and practical examples, increasing grades by 20% on average
- Explained technical concepts to students in the introduction of Microsoft Word, Excel, Access, and PowerPoint, resulting in higher proficiency and overall course performance

## PROJECT

**Legal Document Tool | RAG, FastAPI, LangChain, Ollama, ChromaDB**
- Built a retrieval-augmented generation (RAG) pipeline for semantic search and question answering over legal PDF documents
- Designed document preprocessing pipelines including PDF parsing, text normalization, chunking, and embedding generation using PyMuPDF, spaCy, and sentence-transformers
- Stored and queried document embeddings using ChromaDB to enable efficient vector-based retrieval
- Developed FastAPI endpoints to support document ingestion, indexing, and LLM-based response generation via Ollama

**YouTube Comment Sentiment Analysis | NLP, Python, Pandas, NumPy**
- Collected and processed large-scale YouTube comment data using YouTube Data API v3
- Fine-tuned a RoBERTa-based NLP model for sentiment classification using TensorFlow
- Deployed the project on Streamlit dashboard, allowing users to input a YouTube URL to explore real-time sentiment analysis

**Kaggle - DS&AI Club | NumPy, Pandas, Scikit-learn, Python**
- Conducted exploratory data analysis (EDA) on structured healthcare datasets that have over 250,000 data points and oversampled data for higher accuracy
- Achieved 85% classification accuracy through feature selection and model tuning and developed a Streamlit dashboard to visualize predictions and user inputs