Andrew Chen

510-688-0816 | andrewchenyihan@gmail.com | https://www.linkedin.com/in/andrew-yihanchen/ | https://github.com/Andrewchenyh

PROFESSIONAL SUMMARY

Data-driven Statistics and Economics student at UC Davis with a 3.94 GPA. Skilled in Python, R, SQL, and machine learning, with hands-on experience analyzing large datasets, designing data pipelines, and building predictive models. Strong foundation in statistics and algorithms, supported by research experience and multiple end-to-end projects. Fast learner who enjoys tackling complex problems, refining models, and translating messy data into clear insights.

EDUCATION

University of California, Davis

B.S. in Statistics & B.A. in Economics | Expected Graduation: May 2027

- Cumulative GPA: 3.94 / 4.00
- Awards: Dean's Honor List in Fall 2023, Winter 2024, Spring 2024, Spring 2025
- Certificate: IBM Data Science Certificate, Macquarie University Excel Skills for Business

PROFESSIONAL EXPERIENCE

Operational Management Intern

PAGO LLC | Aug 2025 – Sep 2025

- Analyzed large operational datasets using SQL and Excel to uncover performance trends, automate reporting, and support data-driven decision-making.
- Built a centralized tracking system for 5,000+ SKUs and worker KPIs, streamlining data pipelines and reducing manual processing by 5 hours per week.
- Recommended a revised workflow based on data analysis, which was implemented and led to a 20% reduction in average order processing time.

Student Assistant

UC Davis Technology Transfer Office | Aug 2025 – Present

- Manage weekly intake and initial processing of new Record of Invention submissions, ensuring complete documentation and smooth entry into the university's technology commercialization pipeline.
- Verify sponsor information, MTAs, and supporting documents with high accuracy to maintain compliance and prepare cases for patent evaluation.
- Organize and maintain sensitive legal and technical files for the university's patent portfolio, ensuring accuracy and confidentiality for all case files.

Research Assistant

UC Davis Department of Economics | Sep 2025 – Present

- Collected and digitized 50 years of weekly central bank gold reserve data from The Economist for a historical macroeconomics project studying international monetary transmission during the pre-WWII Gold Standard era.
- Designed an optimization algorithm leveraging each issue's 3-week coverage pattern, reducing required downloads to 34.6% while preserving full data integrity.
- Automated the end-to-end pipeline using Selenium for large-scale PDF scraping and Python tools such as pdf2image and Gemini API to convert scanned documents into structured CSV datasets.

PROJECTS

Fitness Tracker

- Engineered features from raw accelerometer and gyroscope time-series data using low-pass filters, Fourier transform, PCA, and clustering to distinguish exercise patterns across 6 movements.
- Built a visualization pipeline that generated time-series plots for all exercise–participant combinations, comparing accelerometer and gyroscope axes to uncover signal variation patterns.
- Evaluated multiple ML models with grid search and feature-engineering experiments. Tuned a Random Forest model achieving 99.46% classification accuracy, and designed a rep-counting algorithm with a MAE of 0.88 reps.

Audio Sentiment Classification

- Designed and trained a CNN in PyTorch with custom convolutional blocks, early stopping, and a Learning Rate Scheduler to classify 8 audio emotions categories in a fully end-to-end deep learning pipeline.
- Engineered key audio features including MFCCs, Spectral Centroid, Root Mean Square and Zero Crossing Rate from raw audio and implemented audio augmentation to expand training data 5× and improve robustness across audio files.
- Improved generalization through regularization, validation-based tuning, and early stopping strategies, ultimately achieving 93.33% test accuracy and 0.30 test loss on the 8-class emotion classification.

SKILLS

- **Programming:** Proficient in Python and R. Experienced in algorithm design, data pipeline development, web scraping/automation, and building end-to-end workflows.
- **Data Visualization:** Experienced with Python (matplotlib, seaborn, plotly) and R (ggplot2, tidyverse) to build exploratory and analytical visualizations. Skilled in Tableau for dashboards and interactive data storytelling.
- **AI & ML:** Experienced with supervised learning (Random Forest, SVM, KNN, Naive Bayes) and unsupervised learning (PCA, clustering). Skilled in deep learning (CNNs) using Pytorch, feature engineering, model tuning (grid search), and evaluation (accuracy, recall, MAE). Familiar with modern AI tools and APIs to streamline workflows.
- **Data Querying & Analysis:** Strong SQL skills (JOINs, CTEs, window functions) for data extraction; Advanced Excel skills (VLOOKUP, pivot tables, conditional formatting) for clear presentation.
- **Soft Skills:** Self-driven and proactive learner who adapts quickly to new environments; Strong communicator who collaborates effectively across teams; Highly organized with excellent time-management and attention to detail.