

# Thien Toan Vu

## Data Analyst

(714) 417-4049 • [Email](#) • [Website](#) • [GitHub](#) • [LinkedIn](#)

## Education

### University of California, Irvine

September 2021 to June 2024

- ❖ Computer Science, BS • Intelligent Systems • Dean's Honors List (4 quarters) • GPA: 3.233
- ❖ Relevant Coursework: Information Retrieval, Introduction to Data Management, Machine Learning and Data-Mining, and Artificial Intelligence in Biology and Medicine.

### Google Data Analytics Professional Certificate

June 2023 to January 2024

- ❖ Acquired comprehensive skills in data cleaning, organization, and analysis using spreadsheets, SQL, and R programming. Proficient in visualizing data findings using Tableau and other visualization platforms.
- ❖ Gained an immersive understanding in data ethics, integrity, and decision-making within a data analysis framework. Applied problem-solving techniques and conducted case studies, demonstrating a strong foundation in data collection, aggregation, and calculations for effective decision-making processes.

## Work Experience

### Fablix: Movie Database Web App

September 2023 to December 2023

- ❖ Architected and deployed Fablix by setting up AWS infrastructure, MySQL databases, and Tomcat servers. Implemented ETL pipelines for parsing large XML files, ensured secure login with SHA256 hashing and sessions, and incorporated features like full-text search, auto-complete with cache, bot detection using reCAPTCHA, and protection against SQL injection attacks via PreparedStatements, etc.
- ❖ Achieved a 50% boost in website performance through strategic optimizations like MySQL connection pooling, replication, and Apache load balancing, emphasizing a data-driven approach to problem-solving and system optimization.

### Supervise Me: Multi-label Image Classification

March 2023 to June 2023

- ❖ Analyzed DNA damage patterns in mice cells exposed to ionizing radiation, utilizing LeNet-5 architecture and conducting thorough hyperparameter tuning. Achieved top accuracy of 84.78% on the single-label model and 82.7% on the multi-label model.
- ❖ Led end-to-end project planning, encompassing data retrieval from AWS S3, data preprocessing, and model training. Ensured replicability by efficiently organizing the GitHub repository, documenting the research process, and leveraging tools like PyTorch, PyTorch Lightning, AWS boto3 API, and Weights and Biases for streamlined experimentation and analysis in a GPU/CPU environment.

## Skills

- ❖ **Languages:** Python, SQL, R Programming, Java, C++, C, JavaScript, HTML/CSS, C#
- ❖ **Additional Tools/Frameworks:** MySQL Workbench, RStudio, Git, GitHub Actions, RESTful APIs, JMeter

## Other Experience

### Anteater Photography Club (Event Coordinator)

July 2023 to Current

- ❖ Coordinated logistics and enhanced member engagement through strategic planning of meet-ups, venue securing, and fostering collaborative partnerships within the photography community.

### Y Tea Cafe: Boba, Fries, Crawfish (Manager)

May 2019 to September 2021

- ❖ Demonstrated adaptability and strong problem-solving skills through resolving customer issues when arrived while managing daily operations at a bustling restaurant.