Thien Toan Vu

(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: Data Structure Implementation and Analysis, Introduction to Probability and Statistics for Computer Science, Design and Analysis of Algorithms, Information Retrieval, Introduction to Artificial Intelligence, Introduction to Data Management, Machine Learning and Data-Mining, Applications of Probability in Computer Science, Artificial Intelligence in Biology and Medicine, and Introduction to Data Science

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.

Tableau Business Intelligence Analyst Professional Certificate   **January 2024 to Present**

- ❖ Learning Business Intelligence Analytics with expertise in data manipulation, preparation, and analysis using Tableau Public, resulting in the ability to craft compelling data visualizations that reveal patterns and drive actionable insights.
- ❖ Learning skills in presenting and reporting data effectively, with a focus on constructing engaging narratives that enable stakeholders to explore and understand insights, coupled with a strong foundation in data management.

Experience

Case Study: How does a Bike-share Navigate Speedy Success?

- ❖ Conducted a comprehensive data analysis for a bike-sharing service, differentiating between annual members and casual riders by analyzing a dataset spanning 12 months in 2022.
- ❖ Cleaned and processed data using Python libraries, addressing missing values, outliers, and creating new columns for enhanced analysis, resulting in a refined dataset stored in an SQLite database.
- ❖ Uncovered key insights, challenging assumptions such as casual riders having longer average ride durations compared to annual members, identified bike preferences, usage patterns across days of the week, and highlighted significant starting and ending stations for both rider categories.

Fablix: Movie Database Web App   **September 2023 to December 2023**

- ❖ Architected and deployed Fablix by setting up AWS infrastructure, MySQL databases, Tomcat servers, and backend servlets Java files.
- ❖ 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**

- ❖ Focused on multi-label image classification to analyze DNA damage patterns in mice cells exposed to ionizing radiation. Utilized LeNet-5 architecture and conducted rigorous 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.
- ❖ Collaborated with domain experts, including Dr. Lauren Sanders and Dr. Sylvain Costes from NASA, to refine the project's scientific approach. Acknowledged for leading a successful team effort in tackling challenges related to space radiation and contributing valuable insights to the field.

Skills

- ❖ **Languages:** Python, SQL, R Programming, Java, C++, C, JavaScript, HTML/CSS, C#
- ❖ **Tools/Frameworks:** Pandas, Numpy, Seaborn, Matplotlib, MySQL Workbench, RStudio, Tableau Public, VSCode, Google Sheets/ Excel, Git/ GitHub, RESTful APIs

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.

Y Tea Cafe: Boba, Fries, Crawfish (Manager)   **May 2019 to September 2021**

- ❖ Demonstrated adaptability and strong problem-solving skills through resolving customer issues when arriving and managing daily operations