Tai Duc Vu

taivu@stanford.edu • +1 (650) 709-7589
taivu1998.github.io • github.com/taivu1998

Education

Stanford University Stanford, CA

- B.S., Computer Science • GPA: 4.2 / 4.0 • Expected Jun 2022
- **Coursework:** Data Structures & Algorithms • Computer Systems • Multiprocessing & Multithreading • Machine Learning • Deep Learning • Artificial Intelligence • Natural Language Processing • Computer Vision • Information Retrieval • Reinforcement Learning • Web Applications • Android Programming • Game Development • Networking C++ Programming • Scientific Python • Linear Algebra • Probability & Statistics • Calculus • Economics • Physics Cryptography • Scientific Writing

Work Experience

Facebook Menlo Park, CA
Software Engineering Intern *Jun 2020 - Sep 2020*

- Augmented latency profiling tools with module level debug information. Improved their runtime efficiency by 5x.
- Optimized Conv1D and channel shuffle operations with XNNPACK, boosting operator level performance of speech and natural language understanding models on mobile devices by 10x. Used **C++**, **Python**, and **PyTorch**.

TreeHacks Stanford, CA
Software Development Fellow *Jun 2020 - Sep 2020*

- Developed a web application that allowed 2,000 hackers to create profiles, propose project ideas, browse the list of other participants, make connections, explore challenges, filter projects, and form teams. Used **JavaScript**, **React**, **MongoDB**, **Express.js**, **Node.js**, **HTML**, and **Sass**.
- Deployed the website using **Heroku** and **AWS**.

Ronald McDonald House Charities San Diego, CA
Mobile App Developer, Volunteer *Jul 2020 - Sep 2020*

- Built a mobile application in **JavaScript** and **React Native** that provided contact information, check-in, daily events, maps, FAQs, and COVID-19 screening, supporting 130 families daily.

The Stanford Daily Stanford, CA
Web Developer *Apr 2020 - Jul 2020*

- Built front-end components in **TypeScript**, **React**, **HTML**, **CSS**, and **Next.js** to enhance the responsiveness, content organization, and navigation design of article webpages, serving 1,000 readers daily.
- Adopted the **Agile** project management methodology.

Tarjimly Mountain View, CA
Software Developer, Volunteer *Oct 2019 - Dec 2019*

- Implemented programs in **Python**, **Java**, and **SQL** to process data from server and filter 18,000 translators across 32 languages based on the daily requests of 21,000 users.
- Designed machine learning algorithms and pipelines using **Scikit-learn**, **Pandas**, and **MLFlow** to optimize the process of matching users with translators, reducing the average wait time to 2 minutes.

Vietnam Posts and Telecommunications Group Hanoi, Vietnam
Software Engineering Intern *Jun 2019 - Aug 2019*

- Researched image processing techniques and convolutional neural network architectures.
- Developed deep learning and computer vision models to extract texts, numbers, and logos from images of bank cards, achieving 89% accuracy. Used **Python**, **C++**, **TensorFlow**, **NumPy**, and **OpenCV**.
- Embedded the models in iOS and Android apps using **Java** and **Swift**, improving the experience of 40,000 users.

ViCare Corporation Hanoi, Vietnam
Data Analytics Intern *May 2017 - Jul 2017*

- Analyzed daily statistical data of 2 million customers, assisting managers to make sales and marketing decisions.
- Developed tools using **Python**, **NumPy**, **Pandas**, and **Matplotlib** to automate data cleaning, data analysis, and data visualization, speeding up the data analytics pipeline by 10x.

Research Experience

Stanford Machine Learning Group Stanford, CA
Undergraduate Researcher *Apr 2020 - Present*

- Write scripts in **Python**, **NumPy**, and **Pandas** to download and preprocess satellite images and driver labels corresponding to global forest loss events.
- Implement data augmentation techniques. Develop CNN and LSTM models using **PyTorch Lightning** to classify drivers of forest loss from satellite imagery.

Stanford InfoLab Stanford, CA
Undergraduate Researcher *Feb 2020 - Apr 2020*

- Set up an input pipeline for data loading, image preprocessing, and data augmentation.
- Designed Mask R-CNN and U-Net models in **PyTorch** to detect, localize, and track underwater corals in videos.

Computer Science Research Lab Stanford, CA
Member of AI Group *Sep 2019 - Dec 2019*

- Researched model compression techniques, including network pruning and quantization.
- Utilized regularization and pruning to reduce the computational cost of ResNet by 15% with a 2% increase in accuracy. Used **PyTorch Lightning** and **Weights and Biases**.

Projects

IntentBot: Building Machine Learning Systems For Automated Intent Detection

- Built Naive Bayes, Softmax Regression, SVM, LSTM, and BERT models for identifying different intents from user queries, achieving 99% accuracy and 0.99 F1 scores. Used **Python**, **NumPy**, **Pandas**, **PyTorch**, **TensorFlow**, **Scikit-learn**, and **Hugging Face**.

Photo Sharing Web Application

- Built a photo-sharing web application that supported user authentication, user profiles, user listing, photo sharing, favorite lists, commenting, activity feeds, etc. Used **JavaScript**, **React**, **HTML**, **CSS**, **Express.js**, **MongoDB**, and **Node.js**.

GANime: Generating Anime Character Drawings from Sketches

- Developed neural style transfer, Pix2Pix, and CycleGAN models for generating realistic colorized anime characters from sketch drawings, attaining an FID score of 220.5 and an SSIM index of 0.76. Used **Python**, **TensorFlow**, **AWS**, and **Google Colab**.

FlapAI Bird: Training AI Agents to Play Flappy Bird

- Implemented an AI program to play Flappy Bird, achieving scores of 2,000+. Applied reinforcement learning approaches such as SARSA, Q-learning, function approximation, and deep Q networks. Used **Python**, **PyTorch**, **Pygame**, and **OpenAI Gym**.

Stock Charts for Traders

- Built an interactive web application in **Python**, **TypeScript**, and **React** to track and visualize daily movements of stock data, allowing users to devise potential trade strategies.

Mapping Income Distribution with Machine Learning

- Built a machine learning program to predict income distribution in California based on satellite imagery, achieving a mean absolute error of $50/month. Used **PyTorch**, **Scikit-learn**, **Scikit-image**, **NumPy**, and **Pandas**.

Shiptivitas To Do App

- Implemented a to-do list web application based on a kanban board using **JavaScript**, **React**, **Node.js**, **SQLite3**, **HTML**, and **CSS**, which supported freight shippers to manage their shipping requests and increase productivity.

Skills

Languages: Python • C • C++ • JavaScript • TypeScript • Java • SQL • Swift • Assembly • R • MATLAB
Technologies: TensorFlow • PyTorch • Keras • Scikit-learn • AWS • Google Cloud • OpenCV • NLTK • SpaCy NumPy • Pandas • HTML • CSS • React • Node.js • MongoDB • Express.js • Next.js • React Native • Django Sass • Heroku • Spark AR • Git • Linux • Visual Studio Code • XCode • Android Studio

Awards

International Mathematical Olympiad: Bronze Medal
International Mathematics Local Tournament: First Prize
Hanoi Open Mathematics Competition: First Prize
Vietnamese Mathematical Olympiad: First Prize
American Mathematics Competition: First Prize
Regional Mathematics Competition: Gold Medal

Publications

1. Vu, Tai, et al. "How Not to Give a FLOP: Combining Regularization and Pruning for Efficient Inference." ArXiv:2003.13593 [Cs, Stat], Mar. 2020.
2. Vu, Tai, and Tran, Leon. "FlapAI Bird: Training an Agent to Play Flappy Bird Using Reinforcement Learning Techniques." ArXiv:2003.09579 [Cs], Mar. 2020.