Jameson Thai

408-520-8203 | jameson.k.thai@gmail.com | https://www.linkedin.com/in/jamesonthai | https://jamesonthai.github.io

SKILLS
Languages: Python, Java, SQL, PHP, R, JavaScript, C, HTML, CSS, Matlab, XML
Technical Skills: Data Mining, Data Analysis, Data Visualization, Natural Language Processing, Convolutional Neural Nets
Technical Knowledge: Time Series Analysis, KNN, Decision Trees, SVM, Logistic Regression, Clustering, Naïve Bayes
Working Knowledge: Pandas, PyTorch, Keras, Tensorflow, Plotly, Numpy, Matplotlib, Spark, OpenCV, Keras, Scikit-learn, AWS EC2

EXPERIENCE
Directed Research Developer - **University of Southern California**   September 2018 – Present
Technologies Used: Python, Docker, Tensorflow.js, Java, Bash, PostgreSQL, Vagrant, Scikit-learn, Git
- Develop a **cloud-based** framework to study changes over the **development history** of software systems using **quality metrics.**
- Visualize and construct a **Tensorflow.js predictive model** for predicting from a large dataset of git repositories when, how, and to what extent developers influence software quality attributes.

Data Science Intern - Warner Bros. Entertainment   June – August 2019
Technologies Used: Python, DCU API, Keras, TensorFlow, AWS EC2, Scikit-learn, OpenCV, Matplotlib, Numpy, Plotly, Pandas
- Constructed a **Python image recognition system** for identifying Batman in any DC Universe comic page with a **76% AUC**.
- Performed feature engineering, image augmentation, transfer learning, and hyper parameter tuning on AWS EC2 model.
- Visualized layer activations in Xception CNN model in presentation to WB executives, data team and DCU stakeholders
- Automated the collection and parsing of DC **comic images** by deploying a **Python web crawler** calling the DCU API.
- Developed **Python Pipeline** for augmenting image dataset of DC characters across **different eras and art styles** into model.

Software Engineer Intern - San Jose State University   September 2017 – May 2018
- Improved existing **SQL** queries and **PHP** sanitization resulting in faster querying for loading web assets and storing user data.
- Overhauled ticketing system with **Google apps scripts** resulting in greater scalability and task completion in over 60 projects.
- Visualized user attendance and subscription data on performances and genres to gain **actionable insights** for stakeholders.

PROJECTS
Pommerman Competition   August 2019 – Present
Project Repository: https://github.com/Pommerman
Technologies Used: Python, Conda, Gym, Tensorflow, Pytorch, OpenCV, Numpy
- Construct a **Python Pytorch** agent to play the game Bomberman expanding on existing architectures for **multi-agent learning**.
- Collaborated and led eight students to produce **two unique agents** for competing in a **free-for-all** Bommerman game mode.
- Program and explore deep reinforcement learning algorithms such as **PPO** and **A2C** across different reward states.

Riot Champion Trends and Item Recommender   December 2018 – August 2019
Project Repository: https://www.kaggle.com/jamesonthai/can-gold-income-and-damage-trade-determine-wins
Technologies Used: Python, R, Riot API, Scikit-learn, Plotly, ggplot2, Tidyverse, Highcharter
- Automated Python **information retrieval** for **parsing**, **cleaning**, and compiling a **large dataset** of over **400,000** JSON **files**.
- Visualized a user's performance by comparing their match and timeline data to expert and similar players using **R** and **Plotly**.
- **Recommended items** based on **time series data analysis** on average gold income and damage traded with **Python**.

EDUCATION
University of Southern California   Anticipated Graduation December 2020
Master of Science in Computer Science, concentration in Data Science

San Jose State University   Graduated May 2018
Cum Laude (GPA 3.62)
Bachelor of Science in Computer Science, Minor in Mathematics