

Jun-Wei (Jim) Lin

Irvine, CA 92617

✉ junwei1@uci.edu | 🏠 castman.net | 📧 jwlin | 🌐 junwei-lin

Education

University of California, Irvine

PH.D. CANDIDATE IN SOFTWARE ENGINEERING. GPA: 3.96/4.0

Irvine, CA

Sept. 2016 - Present

National Tsing Hua University

M.S. IN COMPUTER SCIENCE. GPA: 4.0/4.0

Hsinchu, Taiwan

Sept. 2006 - Jun. 2008

National Tsing Hua University

B.S. IN COMPUTER SCIENCE. GPA: 3.07/4.0

Hsinchu, Taiwan

Sept. 2002 - Jun. 2006

Skills

Programming	Python (3+ year experience); Java, C++, JavaScript, SQL, Bash (basic familiarity)
Web	Django, jQuery, Bootstrap
Continuous Integration	Jenkins, Robot Framework, Selenium, PhantomJS, JMeter
Machine Learning and NLP	scikit-learn, gensim, NLTK, jieba

Work Experience

Pycone

CO-FOUNDER

Taipei, Taiwan

Jan. 2017 - Present

- Co-founded Pycone, a startup offering online Python courses (2400+ students).
- Designed and lectured the online course "Web Crawling with Python" (750+ students).

University of California, Irvine

GRADUATE STUDENT RESEARCHER

Irvine, CA

Sept. 2016 - Present

- Interested in software analysis and testing. Focusing on leveraging machine learning and natural language processing techniques for mobile/web test automation.
- Published peer-reviewed articles and conference papers with 140+ citations.

QNAP Inc.

SOFTWARE ENGINEERING INTERN

Taipei, Taiwan

Jul. 2016 - Aug. 2016

- Automated manual testing and introduced automatic acceptance/regression testing flow for Q'Center, a centralized platform for managing multiple QNAP NAS.
- Shortened the regression cycle from days to hours.

National Agricultural Library

RESEARCH INTERN

Beltsville, MD

May 2014 - May 2015

- Introduced continuous integration to the software tool development team.
- Designed and developed a queuing system and single sign-on for web services.
- Conducted automated functional, regression and stress testing on web services.

Projects

Predicting Best Answers for Questions on Stack Overflow

CS295 Statistical NLP, Winter 2018

- Applied various ML models (e.g., Random Forest and XGBoost) and NLP techniques (e.g., Latent Semantic Indexing) to predict the best answers for questions on Stack Overflow. Outperformed baseline by 8.5%.
- <https://www.kaggle.com/junwei6/predict-best-answers-using-multifaceted-features>

Kaggle Competition: Rainfall Prediction (7/126, top 6%)

CS273A Machine Learning, Fall 2017

- Used ensembles (e.g., Random Forest and XGBoost) and feature engineering (e.g., missing data handling) to predict rainfall by infrared information.

PTT Web Crawler (230+ Stars and 110+ forks on GitHub)

Feb. 2015 - Present

- A command-line tool to crawl and parse data from PTT, the largest local online community in Taiwan
- <https://github.com/jwlin/ptt-web-crawler>

Bulletin Board for Government Jobs (700+ daily users)

Aug. 2014 - Present

- A website parsing and visualizing open data from Taiwan's government
- <http://opencpa.castman.net>

Awards

- 2018 **SIGSOFT CAPS Travel grant to attend ICSE 2018**, ACM SIGSOFT
- 2016 **Chair's Award and Graduate Dean's Recruitment Fellowship**, University of California, Irvine
- 2014 **Government Fellowship for Studying Abroad**, Ministry of Education, Taiwan

Selected Publications

BOOKS

Web Scraping and Data Analysis with Python (in Chinese)

JUN-WEI LIN AND HUBERT LIN

- Publisher: DrMaster Press. Oct. 2018. ISBN: 9864343386.

CONFERENCE PAPERS

Search-Based Energy Testing of Android

REYHANEH JABBARVAND, JUN-WEI LIN, AND SAM MALEK

- 41th International Conference of Software Engineering (ICSE 2019), Montreal, Canada, May 2019 (21% acceptance rate).

Nemo: Multi-Criteria Test-Suite Minimization with Integer Nonlinear Programming

JUN-WEI LIN, REYHANEH JABBARVAND, JOSHUA GARCIA, AND SAM MALEK

- 40th International Conference of Software Engineering (ICSE 2018), Gothenburg, Sweden, May 2018 (21% acceptance rate).

Using Semantic Similarity in Crawling-Based Web Application Testing

JUN-WEI LIN, FARN WANG, AND PAUL CHU

- 10th IEEE International Conference on Software Testing, Verification and Validation (ICST 2017), Tokyo, Japan, Mar. 2017 (27% acceptance rate).

JOURNAL PAPERS

The i5k Workspace@NAL—enabling genomic data access, visualization and curation of arthropod genomes

MONICA POELCHAU, CHRISTOPHER CHILDERS, GARY MOORE, VIJAYA TSAVATAPALLI, JAY EVANS, CHIEN-YUEH LEE, HAN LIN,

JUN-WEI LIN, AND KEVIN HACKETT

- Nucleic Acids Research, Vol. 43, Issue D1, Jan. 2015

Courses

UCI

INF211 Software Engineering (A+), INF221 Software Architecture (A+), INF225 Information Retrieval (A+), CS273A Machine Learning (A), CS295 Statistical NLP (A) ,