

# Jun-Wei (Jim) Lin

Irvine, CA 92617 • junwel1@uci.edu • <https://github.com/jwlin> • <http://castman.net>

## Education

---

- University of California, Irvine. Ph.D. Candidate in **Software Engineering**. GPA: 3.95/4.0 2016 – present
- National Tsing Hua University, Hsinchu, Taiwan. M.S., **Computer Science**. GPA: 4.0/4.0 2008
- National Tsing Hua University, Hsinchu, Taiwan. B.S., **Computer Science**. GPA: 3.07/4.0 2006

## Technical Skills

---

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

## Work Experience

---

### Co-founder, Pycone Jan. 2017 – present

- Co-founded Pycone, a startup offering online Python courses (1800+ students).
- Designed and lectured an online course “Web Crawling with Python” (630+ students).

### Graduate Student Researcher, University of California, Irvine Sep. 2016 – present

- Focusing on software testing research, especially on test suite reduction and web app testing with machine learning and natural language processing techniques.
- Published peer-reviewed articles and conference papers with 110+ citations.

### Software Engineering Intern, QNAP Inc., Taipei, Taiwan Summer 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.

### Research Intern, National Agricultural Library, Beltsville, MD May 2014 – May 2015

- Initiated and implemented continuous integration for software projects under development.
- Designed and developed the queuing system and single sign-on for a public web service.
- Conducted automated functional, regression and stress testing on web services.

### System Manager, Ministry of Justice, Taipei, Taiwan June 2012 – Sep. 2013

- Managed and executed \$800,000+ IT acquisition projects.
- Developed an online bug tracking system (100+ daily users).

## Projects

---

### Predicting the Best Answers for Questions on Stack Overflow 2018

- Term project of 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 labeled “Python” 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%) 2017

- Term project of CS273A Machine Learning, Fall 2017. Used ensembles (e.g., Random Forest and XGBoost) and feature engineering (e.g., missing data handling) to predict rainfall based on the infrared information.

### PTT Web Crawler (150+ Stars and 80+ forks on GitHub) 2015

- Created a crawler and data parser for the PTT website, the largest local online community in Taiwan (<https://github.com/jwlin/ptt-web-crawler>)

- Parsed and visualized open data from Taiwan's government (<http://opencpa.castman.net>)

---

**Awards**

---

- Chair's Award and Graduate Dean's Recruitment Fellowship (\$7,500), UC Irvine, 2016
- Government Fellowship for Studying Abroad (\$42,000 for three years), Ministry of Education, Taiwan, 2014

---

**Publications** (Google scholar citations: 119) (<https://goo.gl/RpJDax>)

---

**Conference Papers**

- J.-W. Lin, R. Jabbarvand, J. Garcia, and S. Malek, "Nemo: Multi-Criteria Test-Suite Minimization with Integer Nonlinear Programming," International Conference of Software Engineering (ICSE 2018), Gothenburg, Sweden, May 2018. (21% acceptance rate).
- J.-W. Lin, F. Wang, and P. Chu, "Using Semantic Similarity in Crawling-Based Web Application Testing," 2017 IEEE International Conference on Software Testing, Verification and Validation (ICST), Tokyo, Japan, 2017, pp. 138-148 (27% acceptance rate)
- C.-Y. Wu, F. Wang, M.-H. Weng, and J.-W. Lin, "Automated Testing of Web Applications with Text Input," in 2015 IEEE International Conference on Progress in Informatics and Computing, Nanjing, 2015, pp. 343-347.
- J.-W. Lin, C.-Y. Huang, and C.-T. Lin, "Test suite reduction analysis with enhanced tie-breaking techniques," in 4th IEEE International Conference on Management of Innovation and Technology (ICMIT 2008), 2008, pp. 1228-1233.

**Journal Papers**

- M. Poelchau, C. Childers, G. Moore, V. Tsavatapalli, J. Evans, C.-Y. Lee, H. Lin, J.-W. Lin, and K. Hackett, "The i5k Workspace@NAL—enabling genomic data access, visualization and curation of arthropod genomes," Nucleic Acids Research, p. gku983, Oct. 2014.
- J.-W. Lin and C.-Y. Huang, "Analysis of test suite reduction with enhanced tie-breaking techniques," Information and Software Technology, vol. 51, no. 4, pp. 679-690, Apr. 2009.

---

**Courses**

---

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

---

**Language**

---

- Mandarin (fluent), Taiwanese (fluent)