

Huichao Xue

Address:
1000 W. Maude Ave
Sunnyvale, CA 94085

Phone: (559)834-8424
Email: xue.huichao@gmail.com
Homepage: <http://www.xuehuichao.com>

Senior Software Engineer, Machine Learning and Data Mining

Apr, 2017 - Present



Sunnyvale, CA

I simplified the Machine Learning algorithm that discovers 10^6 dream jobs every day – *Jobs You May be Interested In* ([linkedin.com/jobs](https://www.linkedin.com/jobs)).

- Simplified search index queries – -10% latency, +5% job applies.
- Simplified multi-layer ranking algorithm – -5% empty results.
- Turned two search clusters into one: coordinated the cross-team effort to unify *job search* and *jobs you may be interested in*.
- Sherlock team POC: working with four talented engineers to unravel the mysteries behind bad title recommendations.
- Enhanced engineering team's tooling
 - Double # of model launches with a model management tool;
 - Training data backfilling
 - X-ray (under development)

Software Engineer, Tools and Infrastructure

Oct, 2014 - Apr, 2017



Mountain View, CA

I maintained the testing framework for Google's ContentAds frontend web-server. On an average day, my work means 10^6 quicker/stable tests, or 10^2 happier developers.

- 24% faster regression tests – 7 years of computation savings every single day.
- iRegtest – 5 times faster test development time, by keeping the test environment up. Also mentored intern to extend it to Selenium tests.
- I've also implemented various kinds of tests (and random stuffs).
 - Fuzz-tester – a tool that generates unexpected server input from users and/or backend servers. I built an AI scorer for the effectiveness of generated inputs.
 - Load-tester – tools and metrics to calibrate server's performance under heavy traffic. This tool gates our releases.

Software Engineer, for the rest of the world

www.crowdgrammar.com: Proof-read articles quickly, thoroughly and cheaply, by out-sourcing to native US workers. For \$20, you get a 100-page long PhD dissertation proof-read in 3 days, getting at least one correction per sentence.

<https://github.com/xuehuichao/freemind-latex>: Focus on your ideas, and slides shall generate itself. This program compiles mindmaps into L^AT_EX Beamer, and then PDF slides.

Education

Ph.D. in Computer Science, 2008-2015, University of Pittsburgh, PA, USA



- *Advisor*: Prof. Rebecca Hwa GPA: 3.94/4
- *Research Area*: Natural Language Processing, Machine Learning
- *Dissertation*: *Computational Models of Problems with Writing of English as a Second Language (ESL) Learners*
 - Detecting redundancies (e.g. ...I believe that **if**...)
 - Auto-completing tutor feedbacks (demo at www.xuehuichao.com)
 - Detecting preposition errors (e.g. ...need **of**for speed).

B.S. in Computer Science, 2004-2008, Fudan University, Shanghai, China



- Awards: Rank: 3rd/143, GPA: 3.66/4
 - *OOCL Scholarship* – Nov 2006: highest GPA during that academic year
 - *2nd class People's Scholarship* – Nov 2007 and Nov 2005: top 10%.
- *Bachelor Thesis*: Cross-document Coreference Resolution and Named Entity Recognition (supervised by Prof. Yaqian Zhou and Prof. Xuanjing Huang)

Past Internship Experience

Research Intern – Automatic speech grader



Summer 2013

Princeton, NJ(USA)

- Developed regressors to measure learners' grammar sophistication in TOEFL.
- My regressor helped improve 5% in speech grader accuracy.
- Submitted a conference paper.

Software Engineer Intern – Translation table compression



Fall 2010

Mountain View, CA(USA)

- Goal is to reduce phrase table size in Google's translation engine.
- Merged phrase translations that differ only in named entities.

Software Engineer Intern – Ads impression drop investigation



Fall 2007

Shanghai, China

- Goal is to help engineers quickly find reasons of impression drops.
- Built a tool to find correlated events (e.g. impression bump in other campaigns).

Publications

JOURNAL ARTICLES

Martin Birks, Daniel Cole, Stanley P. Y. Fung, and **Huichao Xue** Online algorithms for maximizing weighted throughput of unit jobs with temperature constraints. In *Journal of Combinatorial Optimization* 26(2), pages 237-250, 2013

CONFERENCE PAPERS

Huichao Xue and Rebecca Hwa. Improved correction detection in revised ESL sentences. In *Proceedings of The 52nd Annual Meeting of the Association for Computational Linguistics (ACL short)*, Baltimore, USA, 2014. Association for Computational Linguistics, 2014

Suma Bhat, **Huichao Xue**, and Su-Youn Yoon. Shallow Analysis-Based Assessment of Grammatical Competence for Automated Speech Scoring In *Proceedings of The 52nd Annual Meeting of the Association for Computational Linguistics (ACL)*, Baltimore, USA, 2014. Association for Computational Linguistics, 2014

Huichao Xue and Rebecca Hwa. Redundancy Detection in ESL Writings. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, Gothenburg, Sweden, 2014. Association for Computational Linguistics, 2014

Huichao Xue and Rebecca Hwa. Modeling ESL word choice similarities by representing word intensions and extensions. In *Proceedings of the 24th International Conference on Computational Linguistics (COLING)*, Mumbai, India, 2012. Association for Computational Linguistics, 2012

Martin Birks, Daniel Cole, Stanley P. Y. Fung, and **Huichao Xue**. Online algorithms for maximizing weighted throughput of unit jobs with temperature constraints. In *FAW-AAIM*, pages 319–329, 2011

Huichao Xue and Rebecca Hwa. Syntax-driven machine translation as a model of ESL revision. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING): Posters*, pages 1373–1381. Association for Computational Linguistics, 2010

PATENTS

Huichao Xue and Dhruv Arya and Nadia Fawaz and Liang Zhang. Recommending relevant positions. US Patent Application No. 15/850483, (*filed, pending*), Dec 21, 2017

Dhruv Arya and **Huichao Xue** and Kevin Kao. Recommending relevant positions. US Patent Application No. 15/828915, (*filed, pending*), Dec 1, 2017

Technical Skills

<u>Programming languages:</u>	Python, C/C++, Javascript, Perl, Java.
<u>Scientific computing:</u>	scipy/numpy, Spark, sklearn, Matlab/Octave
<u>Data collection:</u>	Amazon Mechanical Turk, BeautifulSoup
<u>Software development:</u>	GNU Linux, Emacs, Google AppEngine, ionic, angularjs, MySQL/Sqlite, L ^A T _E X, version control softwares
<u>Programming contests:</u>	Before I entered college, I won 1st prize in China’s regional programming contest – top 4 in Hebei Province (NOIP 2003).