Chiao-Lun Cheng
http://www.thechiao.com
chiao@alum.mit.edu

Education

- **Massachusetts Institute of Technology**   **Cambridge, MA**
  Ph.D., Physical Chemistry, *GPA: 5.000*, 2003 - 2008

- **University of California, Berkeley**   **Berkeley, CA**
  B.S. Chemistry, High Honors, *GPA: 3.844*, 2003
  B.A. Molecular and Cell Biology, High Distinction in General Scholarship, 2003
  Physics Minor, 2003

- **University of California, Irvine**   **Irvine, CA**
  Chemistry and Biology Double Major, *GPA: 3.886*, 1999 - 2001

Work Experience

- **Freelance Data Scientist**   **Dec 2015 - present**
  Taipei, Taiwan
  – Built predictive model for a European supermarket chain to optimize workforce scheduling using predicted hourly sales, including automation of data cleaning, feature discovery, and incorporation of external 3rd-party data

  – Set up big data analytics engine for conducting customer segmentation and credit analysis using transactional data

  – Utilized state-of-the-art neural networks for production line defect detection in electronic components

- **Quantitative Trader**   Alluvium Global Research   **Nov 2015 - Oct 2016**
  Taipei, Taiwan
  – Discovered and implemented algorithms for high-frequency trading in international futures markets

- **Chief Technology Officer**   Minus Inc.   **Apr 2014 - Oct 2015**
  Taipei, Taiwan
  – Analyzed passively-collected usage data using both non-parametric techniques (e.g. Random Forest, Gradient Boosting) and parametric techniques in conjunction with bespoke ad-hoc metrics

  – Used in-memory geospatial index and Bloom filters on the JVM to effect spatially-aware, real-time, repetition-less matching for randomchat users

  – Troubleshooted operational issues by rapidly formulating and validating technical hypotheses over entire stack

  – Increased server efficiency by porting peformance-critical infrastructure components from Python to Clojure

- **Partner**   Katong Capital   **Jul 2010 - Apr 2014**
  Hong Kong
  – Discovered and implemented algorithms for intraday trading in international equities futures markets

- **Associate Consultant**   McKinsey and Company,   **Oct 2008 - Jul 2010**
  New York Business Technology Office   **New York, NY**
  – Post-merger management at a major pharmaceutical company, with an emphasis on synergies from merging sales operations

  – Developed a roadmap for upgrading ad sales systems at a broadcast TV company

  – Product development methodology design, development and rollout planning for a fast-growing IT services company

- **Research Assistant**   Massachusetts Institute of Technology   **Nov 2003 - Aug 2008**
  Supervisor: Professor Troy Van Voorhis   **Cambridge, MA**
  – Computational simulations of electron movement in molecules

- **Research Assistant**   University of California, Berkeley   **May 2002 - May 2003**
  Supervisor: Professor Judith Klinman   **Berkeley, CA**

    – Investigations of quantum tunnelling in a soybean enzyme

- **Research Assistant**   University of California, Berkeley   **Sep 2001 - May 2002**
  Supervisor: Professor Paul Bartlett   **Berkeley, CA**

    – Organic synthesis of a digestive enzyme inhibitor

Publications

- Jeremy S. Evans, Chiao-Lun Cheng, and Troy Van Voorhis,
  "Spin-charge separation in molecular wire conductance simulations"
  Phys. Rev. B 78.16, 165108 (2008)

- Chiao-Lun Cheng, Qin Wu, and Troy Van Voorhis,
  "Rydberg energies using excited state density functional theory"
  J. Chem. Phys. 129.12, 124112 (2008)

- Qin Wu, Chiao-Lun Cheng and Troy Van Voorhis,
  "Configuration interaction based on
  constrained density functional theory: A multireference method"
  J. Chem. Phys. 127, 164119 (2007)

- Chiao-Lun Cheng, Jeremy S. Evans, and Troy Van Voorhis,
  "Simulating molecular conductance using real-time density functional theory"
  Phys. Rev. B 74, 155112 (2006)

Toolbox

Natural Languages: English (native), Chinese/Mandarin (native), Taiwanese (fluent)

Software: Emacs, Grep, Sed, Awk, Bash, LaTeX, Mongodb/Tokumx, Cassandra, MySQL, SQLite, Redis, Nginx, COIN-OR/pulp, Mathematica

Computer Languages: Clojure, R, Python, C++11, Fortran 77

Libraries (Clojure): ztellman/aleph, core.async, ring, compojure, CIDER/nrepl

Libraries (R): data.table, plyr, ggplot2, randomforest, mboost, glmnet, lme4, knitr, rstan

Libraries (Python): Pandas, Numpy, scikit-learn, XGBoost, Tensorflow, Keras, Theano, BeautifulSoup, Selenium, gevent, Cython

Technology concepts: JSON, websockets, greenlets, event sourcing, Automatic Differentiation, kd-trees, Bloom filters, SQL, eventual consistency, cache invalidation, Amdahl's Law, sockets/pipes, shared memory, core pinning

Machine Learning concepts: Bias/Variance, Covariance Shrinkage / Factor Modeling, d-separation, Cross-Validation, Bonferroni corrections, regularization (L1/L2), Rao-Blackwellization, ROC curve, Random Forest, Gradient Boosting, Multilevel Regression, Factorization Machines, Stochastic Gradient Descent, Linear Algebra