Feynman Tsing-Yang Liang

feynman.liang@gmail.com
A03 Churchill College, Cambridge, CB3 0DS, UK - +44 (0)7598 415 931 http://github.com/feynmanliang

Education

University of California, Berkeley Berkeley, CA
PhD (Statistics) *Oct 2016– Oct 2021*

- Awarded NPSC fellowship for 5 years and department block grant for 1 year

University of Cambridge Cambridge, UK
MPhil (Machine Learning, Speech, and Language Technology) *Oct 2015– Sept 2016*

- Fully supported by Dartmouth's Reynolds Fellowship and Amherst's Moore Engineering Fellowship

Dartmouth College Hanover, NH
BEng (Electrical Engineering); ***Cumulative GPA: 3.85/4.00*** *Sept 2012– May 2015*

- National level awards: NSF GRFP and Ford Pre-Doctoral Fellowship honorable mentions
- University level awards: Tau Beta Pi inductee (top 5% of class), William Slesnick Prize (best pure mathematics undergraduate research poster)

Amherst College Amherst, MA
BA (Maths, Biochemistry, and Biophysics); ***Cumulative GPA: 3.73/4.00*** *Sept 2010– May 2014*

EdX.org/ Coursera.org/ Udacity.com Online *(Sept 2010– Present)*

- 130+ MOOCs completed; Coursera Hall of Fame (top 25 internationally in number MOOCs completed)

Experience

Toptal Remote
Freelance Software Engineerg *Feb. 2016–Present*

- Member of Toptal network, a community of the top 3% of freelance developers
- Remotely worked on various consumer and enterprise applications, primarily in full stack engineering roles; Key technologies include: React, Angular, D3, Spark, Kafka, AWS

Databricks — Open Source Team San Francisco, CA
Machine Learning Intern *June–Sept 2015*

- Managed development of Apache Spark MLlib, an open source distributed machine learning library library; reviewed code contributions from open-source community; lead Spark MLlib 1.5's QA process including documentation, code test coverage, and performance testing
- Contributed major features to to Spark 1.5 release, including: discrete cosine transformer, association rule mining, distributed PrefixSpan for sequential pattern mining, distributed GMMs, online variational inference for LDA

Dartmouth College — Lorenzo Torresani Hanover, NH
Computer Vision Research Assistant *Sept 2014 – June 2015*

- Built web-based tool to collect fine-grain medical image annotations from doctors; performed baseline experiment using generic image feature extractor and SVM; experimented using (Long et al, *Fully Convolutional Networks for Semantic Segmentation*, CVPR 2015) to segment regions relevant for diagnosis from tumor tissue stain images
- Extended (Lee et al, *Deeply-Supervised Nets*, JMLR 2015) to use image-level labels for supervision of early layers and fine segmentation labels for later layers

Dartmouth College — George Cybenko Hanover, NH
Cybersecurity Research Assistant *Sept 2014 – Mar 2015*

- Improved (Cybenko et al, *Learning Hidden Markov Models using Non-Negative Matrix Factorization* , IEEE Trans. Inf. Theory 2008) by initializing using a SVD low-rank approximation refined using linear programming to be row-stochastic
- Demonstrated applicability of method on Columbia/Kinetic dataset for learning behavior regimes as well as time-series change point detection on MIT Living Lab WiFi dataset

Bridgewater Associates — Core Management Analytics Westport, CT
Technology Associate Intern *June–August 2014*

- Applied graphical models and MCMC to build interpretable models of employee behavior
- Analyzed and presented workplace performance data to non-technical audience using PCA, hierarchical clustering, and PageRank

Google — DoubleClick Search Kirkland, WA
Software Engineer Intern *June–Sept 2013*
- Implemented time series modeling, segmentation, and anomaly detection in online advertising CTR data using ARIMA, GARCH, and HMM modeling
- Developed versioning feature for DoubleClick ad-campaigns using MapReduce (FlumeJava) and NoSQL (BigTable, Spanner)

Microsoft — Office Division Redmond, WA
Software Development Engineer Intern *May–Aug 2012*
- Developed continuous integration and automated build reporting infrastructure for Office 2013

MAStorage Amherst, MA
Founder *Feb 2011–Feb 2012*
- Founded a student summer storage company with >250 annual customers; sold to All College, Inc

Publications, Posters, and Presentations

Yggdrasil: Learning Big Trees *June 2015–Present*
- Journal article submission to ICML 2016, expected to be shipped with Spark 1.7
- Collaboration with Databricks (Joseph Bradley, Xiangrui Meng), MIT (Matei Zaharia et al), and UCLA (Ameet Talwalkar)
- Prototyped, optimized, and productionised a novel distributed decision tree learning algorithm which partitions data by feature rather than by instance

Detecting Paraphrases Using Recursive Autoencoders *Nov 2015*
- Presentation of (Socher, *Dynamic Pooling and Unfolding Recursive Autoencoders for Paraphrase Detection*, NIPS 2011) at Cambridge Machine Learning Group's Research and Communication Club, see www.talks.cam.ac.uk/talk/index/62311

Visualizing Machine Learning Models *Oct 2015*
Large Scale Topic Modeling: Improvements to LDA on Spark *Sept 2015*
Spark 1.5: Association Rules and Sequential Patterns *Sept 2015*
Iterative Learning Control for Pulsatile Biomimetic Blood Flow *Sept 2014–Aug 2015*
- Manuscript in preparation for submission to *Artificial Organs*
- Engineering thesis project; developed an iterative learning controller for producing biomimetic pulsatile blood flow in a cardiopulmonary bypass heart pump

Transplantation of Eigenfunctions on Isospectral Domains *Apr 2015*
- 8 week research project on eigenfunctions of the wave equation, Dirichlet spectrums, and approximate finite-element solutions
- Submitted to Dartmouth's undergraduate poster competition; awarded William Slesnick Prize (first place in pure mathematics)

Teaching

TA for BerkeleyX CS190.1x: Scalable Machine Learning *Sept 2015–Present*
- Developed new MOOC coursework for teaching scalable machine learning using Spark, updated existing coursework to use latest Spark MLlib APIs (e.g. DataFrames, Pipelines)

Mentor for Dartmouth Emerging Engineers *Sept. 2014—May. 2015*
- Mentored first-year engineering students in multivariable calculus, linear algebra, and LTI system theory

Community TA for Berkeley CS169.2x: Engineering Software as a Service *June–Aug 2013*
- Discussed course material and answered student questions in online discussion forum

Other Skills and Interests

Programming Languages: Javascript, Scala, Python, R, Lua, Clojure, C++, MATLAB, Mathematica
Technologies/Frameworks Spark, Hadoop, AWS, Torch7, Caffe, CUDA, Theano, OpenMPI, Anglican, Stan
Other Skills: LaTeX, Eagle Scout, Chinese (fluent)