Jialin Ding

jding09@stanford.edu • www.github.com/jialinding • (314) 705-4098

EDUCATION

Stanford University, Stanford, CA 9/2014 – present

- BS in **Electrical Engineering**, minor in **Economics**, expected 6/2018
- GPA: 3.98
- Coursework: CS 345S: Data-Intensive Systems, CS 145: Databases, CS 168: The Modern Algorithmic Toolbox, CS 265: Randomized Algorithms, CS 229: Machine Learning, CS 231A: Computer Vision, EE 180: Digital Systems Architecture

RESEARCH

Research Assistant, *Future Data Systems Group, Stanford Computer Science* 4/2017 – present

- Researched methods for efficiently extracting insights from high-volume structured data streams
- Implemented algorithmic optimizations for anomaly detection and frequent itemset mining in MacroBase, an analytic monitoring system
- Developed a novel quantile sketch that maintains formal error guarantees by keeping track of higher-order statistical moments
- Collaborated with large Internet company to incorporate MacroBase into their internal telemetry tools

Research Assistant, *Computational Genomics, Stanford Computer Science* Fall 2017

- Improved performance of binary classification task to distinguish between text related to GWAS and unrelated text
- Explored tradeoffs between Naïve Bayes, Logistic Regression, SVMs
- Used domain knowledge to expand the set of features extracted from training examples
- Built web tool for interactive exploration of results

PUBLICATIONS

- Edward Gan, **Jialin Ding**, Kai Sheng Tai, Vatsal Sharan and Peter Bailis. Compact and Mergeable Quantile Sketches using Moments. *SIGMOD*, 2018. Submitted for review.
- Peter Bailis, Edward Gan, Samuel Madden, Deepak Narayanan, Kexin Rong, Sahaana Suri and **Jialin Ding**. MacroBase: Prioritizing Attention in Fast Data. *ACM Transactions on Database Systems*, 2017. Submitted for review.
- Volodymyr Kuleshov, **Jialin Ding**, Braden Hancock, Alexander Ratner, Christopher Re, Serafim Batzoglou and Michael Snyder. A Machine-Compiled Database of Genome-Wide Association Studies. *25th Conference on Intelligent Systems for Molecular Biology (ISMB),* 2017.

PROJECTS

HMM-on-Logs, *CS 345S Research Project* Fall 2016

- Researched fast algorithms based on Hidden Markov Models to detect anomalous sequences in streams of structured log data
- Achieved 10x speedup over a popular Python HMM library (HMMlearn) using sliding window and caching optimizations

Video Stabilization on FPGA, *EE 109 Project* Spring 2017

- Implemented a real-time video stabilization pipeline on an Altera Cyclone V FPGA
- Programmed using Spatial, a DSL for reconfigurable hardware

Chess Vision, *CS 231A Final Project* Spring 2016

- Computer vision program that uses detects a chessboard and the configuration of pieces on the board from an arbitrary image
- Trained a suite of SVMs to classify pieces based on SIFT and HOG features extracted from a manually constructed dataset

INDUSTRY EXPERIENCE

Software Engineer Intern, *Google*, Mountain View, CA Summer 2016

- Worked on Google Safe Browsing within the Privacy & Security team
- Implemented a MapReduce pipeline to integrate Chrome browser incident data into the evaluation of user downloads

Software Engineer Intern, *Thumbtack*, San Francisco, CA Summer 2015

- Performed projects within Search Engine Optimization team to boost organic search traffic
- Implemented backend services in Go for NLP-based automated text generation and content-based recommendation system