

Firas Abuzaid

Curriculum Vitae

353 Serra Mall, Gates 432

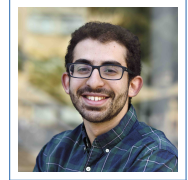
Stanford, CA 94309

+1 (214) 536 4585

+1 (775) 238 3228

fabuzaid@cs.stanford.edu

github.com/fabuzaid21



Education

- 2016–present **Ph.D., Computer Science**, Stanford University
Advisors: Peter Bailis and Matei Zaharia
- 2015–2016 **Ph.D., Course VI-II**, Massachusetts Institute of Technology
Advisors: Sam Madden and Matei Zaharia
◦ Transferred to Stanford in Fall 2016 with advisor
- 2013–2015 **M.S., Computer Science**, Stanford University
Advisor: Christopher Ré
◦ Concentrations: Artificial Intelligence, Information
- 2009–2013 **B.S., Computer Science**, Stanford University
Concentration: Information
◦ Phi Beta Kappa, Tau Beta Pi

Publications

- 2019 F. Abuzaid, P. Kraft, S. Suri, E. Gan, E. Xu, A. Shenoy, A. Anathanarayan, J. Sheu, E. Meijer, X. Wu, J. Naughton, P. Bailis, and M. Zaharia. **DIFF: A Relational Interface for Large-Scale Data Explanation**. VLDB 2019.
- 2019 F. Abuzaid, G. Sethi, P. Bailis, and M. Zaharia **To Index or Not to Index: Optimizing Exact Maximum Inner Product Search**. ICDE 2019.
- 2018 F. Abuzaid, P. Bailis, J. Ding, E. Gan, S. Madden, D. Narayanan, K. Rong, and S. Suri (equal co-authorship). **MacroBase: Prioritizing Attention in Fast Data**. TODS 2018.
- 2018 S. Palkar, F. Abuzaid, P. Bailis, and M. Zaharia. **Filter Before You Parse: Faster Analytics on Raw Data with Sparser**. VLDB 2018.
- 2017 D. Kang, J. Emmons, F. Abuzaid, P. Bailis, and M. Zaharia. **NoScope: Optimizing Neural Network Queries over Video at Scale**. VLDB 2017.
- 2016 F. Abuzaid, J. Bradley, F. Liang, A. Feng, L. Yang, M. Zaharia, and A. Talwalkar. **Yggdrasil: An Optimized System for Training Deep Decision Trees at Scale**. NIPS 2016.
- 2015 S. Hadjis, F. Abuzaid, C. Zhang, and C. Ré. **Caffe con Troll: Shallow Ideas to Speed Up Deep Learning**. SIGMOD 2015, DanaC: Workshop on data analytics at scale.

Work Experience

Research

- 09/2019–present **Research Contractor**, Microsoft Research, Stanford, CA
◦ Continuing work performed as a Research Intern on traffic engineering for wide-area networks
- 06/2019–09/2019 **Research Intern**, Microsoft Research, Redmond, WA
◦ Developed new traffic engineering algorithms for wide-area networks
◦ Mentored by Srikanth Kandula
- 02/2018–11/2018 **Research Contractor**, Facebook, Menlo Park, CA
◦ Developed and deployed Spark pipeline for MacroBase (<https://macrobase.stanford.edu>) to automatically explain anomalies in large-scale datasets for Facebook's Infrastructure organization
◦ Worked with FBlearner team to develop optimization strategies for improving resource efficiency of FBlearner's Inference Platform

- 04/2014–06/2015 **Research Assistant**, *Stanford InfoLab*, Prof. Christopher Ré
- Developed Caffe con Troll, CPU optimizer for Caffe. Resulted in publication
 - <http://github.com/HazyResearch/CaffeConTroll>
 - Worked on multi-round Pregel-like join algorithm for distributed datasets
 - **Technologies and Libraries:** Scala, Spark, SparkSQL, Hive, YARN, Hadoop/HDFS

- 01/2014–03/2014 **Independent Research Project**, *Stanford AI Lab*, Prof. Andrew Ng
- Worked on re-alignment improvements for deep neural networks on speech recognition systems
 - **Technologies and Libraries:** C++, Python, Kaldi, Google Web Speech API

Industry

- 06/2012–09/2012 **Android Engineer Intern**, *Clinkle*, Mountain View, CA
Initial implementation of the Clinkle Android application

- 06/2011–09/2011 **Software Engineer Intern**, *Amazon Lab126*, Cupertino, CA
Developed QuickSettings app for Kindle Fire; debugged, refined UI/UX for other system apps on Android 2.3.3

Teaching

- 2014–2015 **Awards**

- Recipient of the [2014 Centennial Teaching Assistant Award](#)

- 2014–2015 **Instructor**, *Stanford Computer Science Department*

- **CS145**, *Introduction to Databases*, Co-Instructor with Perth Chareernwatttanagul, Summer 2014

- 2014–2015 **Mentor in Teaching Fellow**, *Stanford Computer Science Department*

Responsible for assisting and mentoring new Teaching Assistants for the CS department

- 2013–2015 **Teaching Assistant**, *Stanford Computer Science Department*

- **CS142**, *Web Applications*, Profs. John Ousterhout and Phillip Levis, Spring 2013 - Spring 2014
- **CS145**, *Introduction to Databases*, Profs. Christopher Ré and Jennifer Widom, Fall 2013 - Fall 2014

- 2011–2012 **CS198 Section Leader**, *Stanford Computer Science Department*

Responsible for leading weekly discussion sections to complement lecture for introductory Computer Science classes – CS106A, B, and X

Legal

- 06/2018–present **Expert Consultant for Technical Patent Litigation**, *Shearman & Sterling, LLP*

- **Vendavo, Inc. v. Price f(x) AG et al**, Analyzed and compared Vendavo's and Price f(x)'s respective code bases using state-of-the-art source code comparison techniques. Demonstrated that Price f(x)'s source code was not in violation of Vendavo's copyright protections
- **Additional cases**, Examined patent infringement claims filed against the defendants; for each, identified prior art to invalidate the claims

Other

- 01/2014–01/2015 **Master's Student Liaison**, *Stanford Computer Science Department*

Responsible for communicating and voicing students' feedback and concerns to CS faculty

- 2011–2012 **Academic Theme Associate**, *Stanford Residential Education*

Residential staff position – responsible for creating and planning theme-related programming for the Crothers Global Citizenship dorm

Professional Service

- 2019 **Reviewer**

- NeurIPS (2019), JMLR (2019)

Invited Talks

- 2018 **MacroBase: Prioritizing Attention in Big Data**, *ODSC West 2018*.

- <https://odsc.com/training/portfolio/macrobases-prioritizing-human-attention-in-big-data-2>

- 2018 **Sparsifier: Faster Parsing of Unstructured Data Formats in Apache Spark**, *Spark+AI Summit 2018*.

- <https://databricks.com/session/sparsifier-faster-parsing-of-unstructured-data-formats-in-apache-spark>

- 2017 **MacroBase: Prioritizing Attention in Fast Data**, *HPTS 2017*.

- <http://www.hpts.ws/papers/2017/macrobases-hpts.pdf>

- 2016 **Yggdrasil: Faster Decision Trees Using Column Partitioning In Spark**, *Spark Summit 2016*.
- <https://databricks.com/session/yggdrasil-faster-decision-trees-using-column-partitioning-in-spark>