ALEX KIKUTA CHAVEZ, Ph.D.

AlexKChavez@gmail.com · 734-546-6340

EXECUTIVE SUMMARY

9+ years experience in machine learning software development and data science at both startups and mature organizations, including hands-on-keys implementation of models / algorithms / production software, team management, project management, and ongoing operational responsibilities.

Skills: Software engineering, machine learning, cloud computing, data science, statistics, technical leadership, project management.

Technologies: Python, Scala, R, Java, Amazon Web Services (AWS), Docker, Spark, Hadoop, YARN, Hive, SQL, Vowpal Wabbit, NumPy, SciPy, Linux, Maven, Jenkins, Git, Kubernetes, New Relic, Datadog.

EXPERIENCE

Tech Lead, Research Team, Bridgewater Associates, *Westport, CT* 2/2018 – Present

- Responsible for the software lifecycle of Scala and Python gRPC services for 40+ algorithms, including development workflow, code quality, releases, operations.
- Authored majority of software abstractions for the team, including integration testing frameworks, data loader clients, codec libraries, and tenant configuration support (Scala, Python, gRPC). Authored several analytics and acted as code reviewer for the majority of analytic logic changes.
- Consolidated analytics metadata from disparate sources in a single web service (real-time vs. batch, computation frequency, algorithm specification, QA test cases, change log, documentation).
- Developed and implemented a formal domain model for analytics that simplified a complex codebase and provided stakeholders with a unified vocabulary.
- Developed technical capabilities of team members.

Software Development Manager, Shutterstock, *New York, NY*   9/2015 – 12/2017

- Managed 6 deep learning researchers and software engineers (2017)
    - Core team for eliminating $40M in costs between 2018-2020 via automation of manual review. Opened Boston satellite office and hired 6 team members.
    - Released state-of-the-art face detection and other APIs expected to save $1M/year beginning in 2018.
    - Championed and successfully executed $900K acquisition of NVIDIA DGX appliance as a cost-savings measure to speed up deep learning R&D.
    - Oversaw all aspects of development (containerized microservice architecture, build pipelines, Python, Docker, Kubernetes, Jenkins, AWS).
- Managed 8 applied machine learning engineers and data scientists (2015-2016)
    - Delivered APIs to enhance the customer's search and discovery experience, including [Visually Similar Images](#), [Visually Similar Footage](#), [Reverse Image Search](#), and [Keyword Autotagging](#).
    - Responsible for operations and 99.9% uptime of 17 new and legacy services.
- Machine Learning Engineer, Search Algorithms team (2014-2015)
    - Led development of featurization and ML Scala/Spark library.
    - Trained large-scale L1-regularized linear models on billions of query-image-user features as part of new ranking strategy.
    - Supervised development of low-latency, high-volume spell correction and autocomplete services.
    - Setup ad hoc AWS Spark clusters for data science team members.

Data Science Consultant, *New York, NY*   6/2014 – 1/2015
- Consulted for a security client. Built ML models and wrote papers on findings.

Senior Data Scientist, *Socure, New York, NY*   7/2013 – 5/2014
- Hired and managed two data scientists and three interns.
- Managed AWS identity access, infrastructure, compute, and data pipelines.
- Architected a fault-tolerant, automated fraud prediction pipeline.
- Maximized fraud model performance using gradient boosting, ensembling with client model scores, and feature engineering to demonstrate ROI in terms of fraud dollars saved and manual review costs saved (R, Scala).
- 100% responsible for monitoring strategy, tool selection, and implementation.

- Authored entity resolution and fraud scoring core libraries (Scala, Akka, R).

Data Scientist, *RightAction, Philadelphia, PA*   4/2012 – 7/2013

- Authored pipelines (Scala, Spark, Weka, Vowpal Wabbit, AWS Elastic MapReduce) to train click prediction models on two billion daily impressions extracted from web logs.
- Ad hoc analytics using AWS EMR, AWS Redshift, and R.

Visiting Scholar, *University of Pennsylvania, Philadelphia, PA*   1/2012 – 12/2015

- Published several peer-reviewed journal articles as primary author and statistician.

Test Dev Assoc II, *National Board of Medical Examiners, Phila, PA*   1/2011 – 4/2012

- Conducted original research on detecting cheating and anomalous testing patterns in high-stakes physician licensure exams and published paper as primary author.

Owner, *Divide Zero Networks, Santa Clara, CA*   4/1999 – 8/2005

- Built profitable web hosting company from scratch (PHP, MySQL, HTML).

Research Intern, *Sante Fe Institute, Sante Fe, NM*   5/2004 – 8/2004

- Developed evolutionary game theory models in Java and R.

Operations / IT Intern, *Everypath Incorporated, Santa Clara, CA*   5/1999 – 7/2000

- Customized Bugzilla; managed Redhat Linux server; designed and maintained MySQL database; serviced on-premise and co-located servers.

EDUCATION

- Ph.D., University of Michigan (2010). Psychology, *Formal Modeling, Mathematical Psychology, and Quantitative Methods Specialization.* NSF Graduate Research Fellow.
- M.A., University of Michigan (2007). Statistics.
- B.A., University of Pennsylvania (2005), *Magna cum laude*. Cognitive Science with distinction (minor in Mathematics).