Sourabh Bajaj

http://www.sourabhbajaj.com Mobile: +1-404-594-4787

EDUCATION

Georgia Institute of Technology

Atlanta, GA

Master of Science in Computer Science; GPA: 4.00

Aug. 2012 - Dec. 2013

Email: sourabh@sourabhbajaj.com

Birla Institute of Technology and Science

Pilani, India

Bachelor of Engineering in Electrical and Electronics; GPA: 3.66 (9.15/10.0)

Aug. 2008 - July. 2012

EXPERIENCE

Coursera

Analytics Infrastructure Engineer

Mountain View, CA

Jan 2014 - Present

- **Recommendations**: Service for recommending courses and other entities to personalize learning experience across the platform. Worked on both the offline model training and the online serving stack.
- **Email Notifications**: Developed infrastructure for delivery time optimization and content A/B testing. Built an internal app to run email campaigns for marketing, reactivation and surveys.
- Nostos: Bulk data processing and injection service from Hadoop to Cassandra and provides a thin REST layer on top for serving offline computed data online.
- Dataduct: Open source framework to create and manage data pipelines leveraging reusables patterns to expedite developer productivity. Used for ETL, search indexing and recommendations.
- Data Warehousing: Setup, schema design and management of Amazon Redshift. Built an internal app for access to the data using a web interface. Dataduct integration for daily ETL injection into Redshift.
- Course Dashboards: Backend for Instructor analytics dashboards and learner surveying tools, which helped instructors run their class better by providing data on Assignments and Learner Activity.

Lucena Research Atlanta, GA

 $Data\ Scientist$

Summer 2012 and 2013

- **Portfolio Management**: Created models for portfolio hedging, portfolio optimization and price forecasting. Also creating a strategy backtesting engine used for simulating and backtesting strategies.
- QuantDesk: Python backend for a web application used by hedge fund managers for portfolio management.

Georgia Institute of Technology

Atlanta, CA

Research and Teaching Assistant

Jan 2012 - Dec 2013

- Research Assistant Machine Learning: Research on machine learning for portfolio hedging and replication algorithms. Modeling low-risk & continuous-return strategies. Developed the python library QSTK.
- Teaching Assistant Computational Investing: The online course on Coursera, had more than 100,000 students enrolled. It was featured on the 11 Alive News and the Atlanta Journal Constitution. Involved in creating assignment, exams and conducting recitation sessions. Also taught the on-campus version of the course.

Analog Devices

Bangalore, India

Software Developer Intern

Summer 2011

o 3D Interaction Controller: Prototyped a motion capture system for controlling a 3D image in realtime.

Projects

- QuantSoftware Toolkit: Open source python library for financial data analysis and machine learning for finance.
- Github Visualization: Data Visualization of Git Log data using D3 to analyze project trends over time.
- Recommendation System: Music and Movie recommender systems using collaborative filtering on public datasets.
- Dimensionality Reduction: Comparative analysis of PCA, KPCA and ICA for various classification tasks.
- Mac Setup: Book that gives step by step instructions on setting up developer environment on Mac OS.
- Giraph: Comparative study b/w vanilla Hadoop and Giraph for graph processing using PageRank etc.

Programming Skills

• Languages: Scala, Python, Javascript, C++, SQL, Java Technologies: AWS, Play, React, Flask, Puppet