Achal Shah

Redwood City – CA – 944063
+1 (412) 945 0854 • m@achals.com • github.com/achals
Areas of interest: Scalability, Distributed Systems, Machine Learning

Education

University of Pennsylvania **Phiadelphia**
Master of Science and Engineering, Computer and Information Science *2011–2013*
Relevant Coursework: Internet and Web Systems, Database and Information Systems, Mathematical Statistics, Machine Learning

Institute of Technology, Nirma University **Ahmedabad**
Bachelor of Technology, Computer Engineering *2007–2011*

Experience

Uber **San Francisco, CA**
Sr. Software Engineer *Sept 2015–Current*

- Michelangelo: Currently working on scaling Uber's Machine Learning platform.
- Reliability Platform: Built and operated Hailstorm, an high-throughput, distributed, load generation platform.

Amazon.com **Seattle, WA**
Software Development Engineer *July 2013– Sept 2015*

- Amazon Fulfillment Technologies - Outbound Flow: The team owned services to predict incoming work that warehouses would receive, and the throughput that warehouse associates would achieve, to help optimize planning and operational decisions. I worked on the following:
  - Implemented caching strategies to cut down forecast creation times by 70%.
  - Worked on features to bring the error of both services down under 15%.
  - Contributed to a community-owned monitoring service to automate alarm creation, and drove its adoption across the Fulfillment organization.
- Amazon Fulfillment Technologies - Sortation Team: The team owned software to direct warehouse associates to sort items into individual shipments with minimal human effort. I helped scale the services to serve associates in new regions, and wrote new tools to interact with mechanical conveyance systems.

Raytheon BBN Technologies **Cambridge, MA**
Language Understanding Graduate Intern *May 2012–August 2012*

Worked on the FUSE project, which involved predicting scientific trends based on bibliometric data.

- Implemented feature extraction models for the task of citation prediction. Also implemented cluster quality metrics such as cosine similarity, coherence values, distribution of subjects.
- Prototyped modules to combine semantically related clusters across periods of time.

Academic Papers and Projects

Personality, Gender, and Age in the Language of Social Media **PLOS ONE, September 2013**
HA Schwartz et, al., DOI: 10.1371/journal.pone.0073791

Wrote the prototype to use LDA on a large collection of tweets to cluster words into topic-based word lists. The topics were then correlated with user personality traits.

Distributed Search Engine *CIS 555: Internet and Web Systems* **April 2013**

Built a scalable search engine from scratch. Individual components involved a Mercator-style crawler that started with a list of seed URLs, a component to create an inverted index from the crawled pages, a hadoop implementation to compute the page rank of pages and a servlet-based webserver to serve the results of web queries.

Predicting Amazon Ratings *CIS 520: Machine Learning* **August 2011**

Implemented and compared different machine learning algorithms along with some feature selection methods on a dataset of Amazon product reviews and their corresponding ratings to predict the ratings for unseen reviews.

Automatic Text Summarization *CIS 530: Computation Linguistics* **November 2011**

Built a system to use Named Entity Recognition, co?reference resolution and other lexical features to generate summaries from a collection of documents.

Independent Study: Dependency Eigenwords *with Prof. Lyle Ungar* **September 2012**

Use dependency parse contexts to create eigenwords.

Skills

Fluent in: Java, Python
Familiar with: C/C++, Rust, Javascript, Scala, Haskell
Frameworks & Toolkits: Spring, Matlab, Mallet, Stanford NLP
Operating Systems: Linux/Unix, OS X, Windows
Version Control Systems: Git
Databases and Data stores: Oracle, MySQL, BDB
Web Technologies: Apache Tomcat, Jquery