Ankit Soni

github.com/Ankit1010 linkedin.com/in/ankitsoni9 ankitsoni9@gmail.com (412)251-1593 Palo Alto

EDUCATION

Carnegie Mellon University: B.S. in Computer Science, Math Minor — **June 2014**

EXPERIENCE

Applovin — *Software Engineer – Platform Team* — **Oct 14 – Oct 15**

- Setup and managed data processing infrastructure – including event ingestion to HDFS, Spark cluster & jobserver tuning & management, and a datasources API.
- Used to ingest 30B events daily into 1TB parquet dumps, query tables via JDBC, and powered daily machine learning jobs, business analytics, system debugging.
- Made open-source contributions to spark-jobserver to allow multi-user and HA (via multiple JVMs, akka-cluster, akka-remote)
- Scala, Java, Akka, concurrency, performance & memory optimization, Zookeeper.

Amazon — *Software Development Engineer Intern* — **Summer 2013**

- Built a system to generate product recommendations for customers as they browse the web, using aggregated customer browsing data (Java, Spring, Guava, Python).
- Performed basic data analysis, and tuned an internal machine learning framework to build a data processing pipeline, with incremental updates.
- Analyzed data using Elastic Mapreduce(EMR), deployed on EC2 and stored output on S3.

Avere Systems — *Engineering Intern* — **Summer 2012**

- Built an automated cross-browser/operating-system web-testing framework (Python, Selenium, jQuery)
- Developed system to automatically generate testing code applicable to any browser/OS from a DSL description of the test.
- Wrote a web application (Django) to dispatch tests, and to display, search and filter results of tests. Used to comprehensively test 30+ webpages automatically across 20 platform/browser combinations.

Lunar X Prize CMU Team (under Red Whittaker – CEO Astrobotic Tech) — **Spring 2011**

Publicity Team – Astrobotic's Interactive Lab

- Built an interactive lab that allows users to control a live camera, and other devices through the web. Built a queue for concurrent camera control requests without requiring users to create accounts or login.
- Set up networking stack, NATting, Dynamic DNS, and built both the client-side and server-side of the interactive lab web application. (Ruby-on-Rails/jQuery)
- Resulted in the team getting the most fans among the other competing teams for the Lunar X Prize (228 vs 125 for the 2nd place, +200 fans after project).

PROJECTS

Chess Parser

- Developed a PGN parser to parse chess games into move trees. (Scala, scalaz, atto)
- Adding ability to construct a game tree, and generate training puzzles.

Hadoop/HDFS clone

- Built a simplified clone of Apache Hadoop Map-Reduce and the underlying HDFS
- Users can configure a cluster, and launch map-reduce jobs via a Java API
- Fault-tolerant system – Participant nodes can die & rejoin at any point.Replication factor is maintained, jobs are transparently re-scheduled.

Question Answering System

- Built a system to automatically answer questions based off a corpus of Wikipedia articles (Scala, Stanford Parser, OpenNLP, SistaNLP)
- Built and used components for Information Retrieval (Okapi BM25), Keyword Extraction, Named Entity Recognition, Co-Reference resolution.
- Ranked 2nd out of 15 teams, including graduate student teams.

LiveWeb

- Developed a Chrome extension that allows multiple people to browse and annotate the web together in real-time (Scala, JS, WebSockets)
- Won 3rd place at a Pittsburgh regional hackathon.

Proxy Server

- Built a multithreaded proxy server in C that handles HTTP/1.0 GET requests and binary data (e.g pictures, videos, audio). Supports concurrent requests & caching.

SKILLS

Programming: Java, Scala, Python, SML, Ruby, C, Haskell, jQuery, Bash.
Familiar with Unix command line / shell scripting, git, vim, zsh.
Distributed Systems

EXTRA CURRICULARS

Member of the Chess Club, Plaid Parliament of Pwning (Security Research Group)
SIFE – Active in promoting entrepreneurship on campus via competitions