Harsh Pankaj Panchal

www.harshpanchal.com | panchalhp@gmail.com | +1-607-280-3831 | San Mateo, California

EDUCATION

Cornell University, College of Engineering, Ithaca, NY *Aug 2012 – May 2013*
Master of Engineering in Computer Science, **GPA: 3.93/4.00**

University of Mumbai, Sardar Patel Institute of Technology, Mumbai, India *Jun 2008 – May 2012*
Bachelor of Engineering in Computer Science, **GPA: 3.98 /4.00**

TECHNICAL SKILLS

Languages: Scala, Go, PHP, Python, Java
Datastores: PostgreSQL, DynamoDB, Elasticsearch
Frameworks: Spark, Akka, Play

WORK EXPERIENCE

Software Engineer, Thumbtack Inc. (San Francisco, CA) *Nov 2015 - Present*

- **Aggregating click signal data to use in ranking search results**
  - Built an offline pipeline to extract click signals from relevant events data and made it available for offline analysis and online serving
  - Built the pipeline using Scala and Spark that generated daily snapshots of engagement signals (views, clicks, contacts) by aggregating over millions of events over the past few months
  - Added monitoring dashboards and alerting to validate the generated data to ensure sanity and to make the data immune to upstream changes
- **Reducing latency of search results**
  - Reduced e2e latency of the service that searches and ranks professionals using a customer request by 50%
  - Added relevant monitoring and logging to gain better insight into the slow components in the service
  - Parallelized independent I/O calls, optimized datastore performance by reducing the data size by 70% making it more efficient
- **Building Thumbtack's machine learning platform**
  - Built the entire of Thumbtack's machine learning infrastructure from the ground up that went on to support 10+ models in production including regression models, random forest models and boosted trees
  - Enabled various teams to train, evaluate, serve and monitor their models that powered some of Thumbtack's core product experiences using infrastructure built in Scala and Spark

Member of Technical Staff, Oracle America Inc. (Redwood Shores, CA) *July 2013 – Nov 2015*

- Worked with Oracle's Automatic Storage Management (ASM), which is a volume manager and file system for Oracle's database files, Oracle Exadata storage and Oracle Real Application Clusters
- **Version compatibility between Oracle RDBMS and ASM**
  - Developed a framework to identify the compatibility between different versions of ASM and RDBMS by checking the information about the various dependant patch-set units applied to both components at startup

PROJECT EXPERIENCE

Distributed Key Value Store *Jan 2013 – May 2013*

- Implemented a Key Value Store in Java with Virtual Server Clusters having 1-resilience
- Cluster membership was managed in Amazon's SimpleDB and updates were shared between clusters using anti-entropy

Evaluating page ranks from a web graph *Jan 2013 – May 2013*

- Evaluated page rank of each page given a web graph of 680,000 nodes and 7 million edges using MapReduce
- Computed page rank using simple and block computations and compared the efficiency of both
- Project was implemented in Java and utilized Amazon Elastic MapReduce Web Services