

EDUCATION

M.S., Computer Science	Arizona State University	GPA: 3.67/4	Jan 2016 – Dec 2017
B.Tech., Computer Engineering	IIITD&M, Chennai, India	GPA: 8.15/10	Aug 2011 – May 2015

Relevant Courses

- Distributed Database Systems, Database Management System Implementation, Statistical Machine Learning, Cloud Computing, Software Design, Foundations of Algorithms, Data Mining, Artificial Intelligence

TECHNICAL SKILLS

Programming Languages:	Proficient: Java/Java EE, Python Prior Exp: C++, MATLAB, XML, JavaScript, JSP, HTML, CSS, Scala
Distributed/Big Data Technologies:	Spark, MapReduce, Hadoop, Pig
Cloud Services:	AWS, Google Cloud Platform
Databases:	MongoDB, MySQL, PostgreSQL, HBase
OS, Frameworks & Tools:	Linux/Windows, SPRING, Kafka, Zookeeper, JUnit, Hibernate, Git, Maven, Eclipse

EMPLOYMENT

Senior Software Engineer	Cloudera, Inc. (Palo Alto, CA)	Feb 2018 - Present
---------------------------------	---------------------------------------	---------------------------

- Developing as part of the HBase team (7 members) to [contribute](#) to the Open Source Apache HBase.
- HBase is a distributed, scalable, big data store based on Hadoop as filesystem & Zookeeper as coordination service.
- Attained the position of Apache HBase Committer and PMC within a span of less than 2 years.
- Worked on building and releasing a version of HBase with Java 11 support.
- Worked on Space Quotas Management feature in HBase and fixed several bugs in the feature.
- Worked on the Meta-Metrics feature in HBase and thoroughly tested and added test cases to validate.
- Handle HBase releases upstream and backports of issues to CDH.
- Worked as a “gatekeeper” for CDH 5.16 version which went GA recently.
- Building bash scripts and modifying test suites for Java programs using JUnit framework to increase code coverage

Software Intern (Operations Team)	CYR3CON, Phoenix	May 2017 – Nov 2017
--	-------------------------	----------------------------

Technologies Used: **Python, Hadoop, Map-Reduce, Spark, MongoDB, PostgreSQL, PyCharm, AWS, Linux**

- Delivered results for 40% improvement in data access speed by migrating infrastructure from Postgres to MongoDB.
- Re-designed MongoDB schema by reducing collection & field counts leading to 5% quicker response to REST API endpoints.
- Involved in incremental development with cross-functional teams & deployment to production environment using Git repos.
- Decreased Postgres data transfer time from 40 mins to 7.5 mins through Spark and multithreaded Python implementations.
- Initiated the decision to use Mongo-Engine as object data mapper which now handles 50 GB of data on daily basis.

Software Engineer	Infoview Technologies Pvt Ltd., India	Jul 2015 - Dec 2015
--------------------------	--	----------------------------

Technologies Used: **Java, AWS, Rest/Soap API, SQL, Agile, Kafka, TDD, Eclipse, Windows**

- Reduced SLOC by 1.5k lines by improving object-oriented designs of Java Spring MVC controllers and business logic layers.
- Developed SOAP/REST based APIs using JAX-WS, JAX-RS for logging 0.5 GB of error messages per day to internal GUI.
- Created 3-tier auto-scaling JAVA web service on AWS using EC2, S3 & SQS to dynamically improve concurrency by 10 times.
- Organized team of 7 members to analyze issues of cleaning, de-duplication & cron jobs on 20 GB data pipeline.

PROJECT EXPERIENCE

Foodie Friend (Python, Google App Engine)	Jan 2017 – May 2017
--	----------------------------

- Innovated an automatic scaling Python app that translates menu cards into English, displays pictures & ingredients of dishes and suggests must-try local foods using Google App Engine’s Pub/Sub, Compute Engine, VISION API & BigQuery.

Lyrics based song recommender system (Python, SVM)	Jan 2017 – May 2017
---	----------------------------

- Developed a Multinomial, SVM based song classifier & recommender system which performed with 81.6% accuracy.
- Generated 1500 full English lyrics of mood-based songs from lyricwikia as training data, harnessing Python web parsers.

Hot Spot Analysis (Java, Hadoop, Map-Reduce, Spark)	Aug 2016 – Dec 2016
--	----------------------------

- Accomplished identification of top 50 statistically significant hot spots in Yellow Cab trip data of 1.8 GB by implementing Java Map-Reduce programs in Hadoop & Spark in a team of 5 which led to 100% accuracy in 2.5 seconds.

Runners up in ACM ICPC Multi Provincial Coding Contest held at Chennai, India.	October 2013
---	---------------------