

ANUJ H. SHAH

(213-572-7769) | anujsha@usc.edu | www.anujshah.com | <https://www.linkedin.com/in/anujshah>

EDUCATION

University of Southern California, Los Angeles, CA Masters in Computer Science <i>Natural language processing, Content Detection and Analysis for Big Data, Information retrieval</i>	Aug '15 – May '17 GPA: 3.66/4.0
Mukesh Patel School of Technology Management and Engineering, Mumbai, India Bachelors of Technology in Computer Engineering	Aug '09 – April '13 GPA: 3.71/4.0

EXPERIENCES

- **PHP Intern, Dun & Bradstreet, Malibu, US (Ongoing)** Jan '17 to May '17
 - Working as a PHP developer on RESTful APIs using slim framework
- **Software Engineering Intern, Rubicon Project, Los Angeles, US** Jun '16 to Dec '16
 - Developed web application using PERL and CGI
 - Manifested automation of data flow of ad(s) requests of specific type in/out of Rubicon Project.
 - Enables various teams troubleshoot problems, reducing manual intervention by ~90%.
 - Designed and developed a bash script
 - Toggles configuration for Memcache and Nginx on user space/server
 - Facilitates Memcache and Nginx implementation on multiple servers.
- **Software Engineering Analyst, Accenture (for BestBuy)** May '13 to May '15
 - Developed a PL/SQL script which resolved the issue of duplicate memberships fixing 300K customers.
 - Optimized ~25 queries by using explain plan and adding hints, reducing the result time by avg. 45%
 - Worked on scripts and daemons handling database and files across vendors using C, C++ and PL/SQL.
 - Developed a shell script which automates reporting the concerned team in case of a program(s) failure.
- **Founder, Ninthbox (ecommerce Platform)** Nov '14 to July '15
 - Developed an ecommerce website using open source platform – Opencart, PHP and MySQL.
 - Sole proprietor of ninthbox for individuals to showcase their talent of creativity featuring 1500+ products.
 - Managed all the modules of the website (development, manufacturing, shipping, support, etc.).
 - The website had ~7000 views in span of five months. <http://ecomm.ninthbox.com/>

PROJECTS

- **Context Recognizer framework for developing chat bots** December '16
 - Developed a framework which acts as a core for chat bots, used Reddit's 3M posts containing 7 categories.
 - Used Perceptron, Naïve Bayes, Stochastic Gradient Descent and Support Vector Machine to classify data.
 - Accuracy of ~74% and precision of ~73% was achieved.
- **Search engine for web pages** December '16
 - Built a search engine with Auto complete and spell correction using crawled pages with crawler4j
 - Used Lucene and Pagerank algorithm to rank the result web pages and Apache Solr to index pages.
- **MIME Diversity Analysis** March '16
 - Analyzed MIME diversity of National Institutes of Standards and Technology (NIST) and Text Retrieval Conference (TREC) Polar Dynamic Domain dataset.
 - Executed Byte Frequency Analysis, Byte Frequency Correlation, Byte Frequency Cross Correlation and File header trailer algorithms on ~60GB of dataset.
 - Rendered data visually on D3 graphs. The analysis helped identifying magic bytes in the specific MIME Types which was further contributed to Apache Tika.

TECHNICAL QUALIFICATION

- Languages : JAVA, Python, C, C++, Shell scripting, PERL, PL/SQL
- Web Technologies : HTML, HTML5, CSS, PHP, JAVASCRIPT, JSON, XML, DOM, Bootstrap, jQuery, AJAX
- Databases : MySQL, MS Access, Oracle, Oracle Forms, Oracle Reports
- Tools : SVN, Jira, CVS, git, github, Apache Solr, Apache Lucene