

ROHITH YERAVOTHULA

<https://www.linkedin.com/in/rohith-yeravothula/>

Los Angeles CA 90007

rohith.iitj@gmail.com (323) 620-7351

EDUCATION

- **University of Southern California** **Los Angeles, CA**
Master of Science in Computer Science May 2019
- **Indian Institute of Technology** **Jodhpur, India**
Bachelor of Technology in Computer Science and Engineering, CGPA **8.77/10** May 2015

SKILLS

- **Languages:** Python, Scala, Java, C, C++, Shell scripting
- **Software:** Apache Spark, Kafka, Flink, Docker, Git
- **Cloud Platforms:** AWS, Azure, Google Cloud Engine
- **Web Technologies:** JavaScript, AngularJS2, D3js, HTML
- **Databases:** Cassandra, MongoDB, PostgreSQL, Redis
- **Frameworks:** Flask, Play-Scala, Scala-Spray, Akka-Actors

EXPERIENCE

Research Student, Information Sciences Institute USC **Los Angeles**
Causal Knowledge Organization August 2017 - present

- Built a text annotation system to generate ontological mappings and relationships using **Python, Spacy**, and Brat

Senior Software Development Engineer, Sigmoid **Bangalore, India**
Domain Specific Text Knowledge Graph May 2015 – June 2017

- Developed an Information Extraction system to convert raw text into knowledge graph on **Spark** and **Cassandra**
- Remodeled graph construction algorithm to avoid super nodes which reduced average query time by **20%**
- Created knowledge graph **CRUD REST API**, Web interface to train, test and integrate with external applications

Time Series Real-Time Predictive Maintenance System

- Implemented an unsupervised machine learning algorithm for time series data on **Spark, HDFS**, and **MongoDB**
- Engineered a real-time prediction model on **Spark Streaming** and **Kafka**, that scaled the performance by **100X**
- Collaborated in a team of **3** and developed a web interface to monitor and train the system

Mentored **4 interns** to implement a web scraping tool for e-commerce websites on Apache Spark and MongoDB

Developed ETL pipeline on Apache Spark, HDFS, and AWS to handle **150TB** of data with a response time of **10 sec/query**

Summer Intern, University of British Columbia **Vancouver, Canada**
Cyber Security for Power Grids May 2014 – June 2014

- Designed a mathematical model to represent UBC's power grid as SCADA architecture on I2Sim software
- Improved model's prediction accuracy to **80%** in real-time scenarios by incorporating attack graph concept

Summer Intern, Defense Research and Development Organization **Hyderabad, India**
C Compiler Test Suit May 2013 – June 2013

- Designed and Implemented a C compiler test suit using differential algorithm, CSmith, Python and Shell Scripting

PROJECTS

Sentimental Analysis on Twitter stream May 2017 – June 2017

- Created a real-time distributed sentimental analysis pipeline on **Spark Streaming** and **Stanford NLP** for Twitter data

Autonomous Asset Management System Sep 2016 – Jan 2017

- Coordinated and led a team of **3** to implement a real-time infrastructure management system to monitor **3M** assets
- Eliminated manual monitoring of assets by introducing rule-based mining engine to identify anomalies

Apache Log Analyzer (Hackathon) July 2016 – Aug 2016

- Formulated a real-time apache log analyzer pipeline to provide statistical insights from log messages
- Collaborated with a team of **4** to implement the pipeline on Kafka, Spark Streaming, Elastic search and Kibana

Hybrid Web Scraper May 2015 – Dec 2015

- Designed and built a configurable k-level hybrid Web scraper that extracts documents and loads into MongoDB
- Increased speed to **1.31** folds by re-designing and implementing as distributed web scraper in Spark and Scala

EXTRACURRICULAR ACTIVITIES

- Member and frequent speaker at Sigmoid meet up group consisting more than **4K** developers, Bangalore
- Elected as students' senator and member of students' general body committee at IIT Jodhpur