

## SUMMARY OF QUALIFICATIONS

- Strong experience in applying ML end-to-end from defining the problem, collecting/crowdsourcing the required data, research on state-of-the-art techniques/models, and to production ready systems consumable by the user.
- Specialization in NLP, deep learning, chatbots, working with large amounts of unstructured data and infusing AI in products.
- Innate ability to lead, simplify and communicate clearly with team members, stakeholders, and end users.
- Technical: Python, SQL, Hive, Spark, Tableau, ML frameworks (Sklearn, Tensorflow, Keras)

## EDUCATION

University of Washington, Information School, Seattle, WA (3.93/4.0) June '17  
*Master of Science in Information Management -- Specialization in Data Science & Analytics. Relevant Coursework: Data Science, Machine Learning & Econometrics, Advanced Data Visualization, Information Architecture, Data Scaling, Applied Social Statistics.*

Osmania University, Hyderabad, India (86.85%) June '14  
*B.S. Computer Science and Engineering -- Relevant Coursework: Data Structures, Software Engineering, Database Management Systems, Artificial Intelligence, Data Mining, Information Security, Information Storage and Management*

## WORK EXPERIENCE

SAP Labs, Innovation Center, Palo Alto, CA – Data Scientist II March '17 – Present

- Research and development of state-of-the-art natural language models for intent recognition, entity extraction, and downstream NLU tasks in the context of enterprise/consumer grade chatbots.
- Closed domain Q&A, comprehension and knowledge representation research on structured/unstructured data and documents.
- Partner with engineering on scaling the built models in a production environment.
- I also directly lead all efforts on building a reusable and scalable Data Science pipeline that enables better research, data-driven decision making on algorithms to deliver in production and effective communication within the cross-functional global team.

Concur, Seattle, WA – Product DS Intern June '16 – March '17

- Worked in the Product Intelligence team to build models for customer churn to proactively identify and mitigate attrition risks.
- Informed product strategy for the development of a new simplified Analytics Product for Concur Customers to replace an existing 150M ARR reporting product using NLP/ML.
- Built and automated end-to-end data pipeline for data extraction, preprocessing, mining, prediction and data visualization.

Google Inc., Hyderabad, India – Associate Account Strategist May '14 – Aug '15

- Provided strategic advice and data-driven insights to help advertisers maximize ROI by working closely in a product support role.
- Worked in the Tools team to develop advanced analytics dashboards for the Business Unit to drive business decision making.

Google Inc., Hyderabad, India – SMB Intern May '13 – July '13

- Impact measurement of AdWords campaign optimizations performed by account optimizers.
- Consolidated and analyzed market-wide data to suggest new optimization opportunities and process improvements.

## ACADEMIC RESEARCH PROJECTS

Data for Economic Development - Improving access to Financial Institutions based on Mobile Cellular Data Records in Ghana

- Worked as a Research Assistant for Prof. Joshua Blumenstock to crunch terabyte scale Call Detail Records data using Spark.
- Cleaning, Anonymizing, and Visualizing data for enabling a clear understanding of the social and economic behavior of population.

HCUP (Healthcare Cost and Utilization Project) - Fixed Effects Econometrics ([Project Link](#))

- Worked with Prof. Benjamin Althouse on modeling relationships between vaccination uptake and hospitalization for WA state.
- Built fixed effects model for gauging the impact of vaccination within counties normalizing for demographic characteristics.

Wikitrends - Wikipedia Trend Analysis During US Presidential Elections ([Project Link](#))

- Understanding relationships between Wikipedia and general interest on Google Trends over time for presidential elections.
- Mined extremely large messy data in the form of Wikipedia log files using GraphLab Dato Distributed API on 20 node cluster.