

Research interests : systems (distributed, data-intensive computation), storage and ML

## EDUCATION

---

- **PES Institute of Technology** 2013 – 2017  
*Bachelor of Engineering in Computer Science; CGPA: 9.00/10.0 (Top 5%)*  
*Thesis title: A low-cost method to create chalk-talk videos (supervised by Dr. Viraj Kumar)*
- **Chetan Pre-University Science College** 2011 – 2013  
*Board of Intermediate Education (12th grade); AIR: 340 (Top 1%)*

## PROFESSIONAL EXPERIENCE

---

- **Swiggy** Bangalore, India  
*Senior Machine Learning Engineer* May 2019 - Present
  - Built end-to-end platform to deploy deep learning models at scale using tensorflow serving. Achieved latencies of 15ms at peak 10k RPS using sidecar containers and gRPC network calls
  - Maintaining and operating ML platform with over 5k feature jobs and ~150 scala models in production
  - Streamlined visibility and monitoring of ML models by feature pipeline quality checks, alerts on data drifts, erroneous models, and infrastructure failures. Reduced TAT for model failures from 1 day to 2 hours
  - Developed real-time and end-of-ride map-matching algorithm to snap driver GPS pings to the underlying road network. Powering accurate driver payouts, missing road detection, ETA predictions, and order assignments
  - Architected map-reduce style spatial querying and data manipulation engine for massive storage of point, line, and polygon data. Provides a real-time, cost-effective, and performant solution for efficient OLAP queries
  - Built a low-throughput, high-latency prototyping platform to serve python-based models at scale. Enables data scientists to conduct experiments and validate hypothesis without rewrite in high-performant scala/tensorflow
  - Deployed smart payments model into checkout springboot service with a peak throughput of 2mn requests per day. Generated real-time features from the order flow and wrote JUnit cases for compile-time testing
  - Built performance testing framework to measure how fast models can fetch features and produce results when deployed in critical software systems. Helps gauge whether model is production-ready and within latency budgets
  - Worked closely with data scientists to onboard multi-objective models optimizing for competing metrics – UE, CX
  - Involved in design and solutioning of platform's automated retraining capability, model orchestration framework, centralized feature store, and monitoring and alerting framework
- **Freshworks** Chennai, India  
*Machine Learning Engineer* Oct 2018 - May 2019
  - Migrated legacy codebase from in-memory redis cluster to disk-based cassandra, reducing burn-rate by \$250k per year. Implemented memcached to increase the key-fetch rate further and minimize latency
  - Architected database model for storing normalized term frequency and document frequency across articles, achieving average  $O(1)$  read and write speeds
  - Built APIs for exposing tf-idf ranking model to end customers through a chatbot. Also integrated diverse use-cases like smalltalk, open-domain question answering, gibberish detector, and custom intent detection engine
  - Worked on language-agnostic spell-correct microservice achieving average search complexity of  $O(1)$ , at the cost of pre-calculation time and storage space of  $n$  deletions
- **Noodle.ai** Bangalore, India  
*Associate AI Engineer* Jul 2017 - Sep 2018
  - Built and orchestrated demand forecasting ecosystem for real-time consumption (using R). Wrote DAGs using airflow as the workflow schedule system to run batch jobs
  - Worked on scaling compute by employing SPMD on N cores using parallel backends like doSnow, doParallel in R
  - Developed an incremental learning framework using the global-local ensemble model, where global serves as a long term model and local serves as a short term model
  - Worked on a proprietary ensemble modeling technique consisting of multiple models such as arima, xgboost, croston, prophet to capture the heterogeneity of various time-series
- *Software Engineer Intern* Jan 2017 - Jun 2017
  - Developed an in-house EDA tool which automatically munges data to plot the features, performs statistical tests, and summarize the result
  - Built automated training and prediction backend batch jobs

## RESEARCH EXPERIENCE

---

- **National Institute of Rural Development**

Aug 2016 - Nov 2016

*Research Intern (Remote)*

*Non-invasive blood screening of hemoglobin using smartphone camera (supervisor: Dr. N V Madhuri)*

- Performed chromatic analysis of blood using a smartphone camera and LED light by measuring absorption properties at various wavelengths, using Beer-Lambert's law
- Extracted peak and trough intensity of each pulse and a combination of high-pass filter and fast-fourier transform is applied on the waveform to obtain RGB time-series
- Conducted extensive user-study of 320 participants and deployed app across 60 health camps in rural parts of India

- **Center for Cloud Computing and Big Data**

Jan 2016 - May 2016

*Research Associate*

*Standardization of docker events for auditing in OpenStack environment (supervisor: Dr. Dinkar Sitaram)*

- Extended cloud audit data federation (CADF) specification to support docker and contributed to pyCADF, open-source python implementation of the auditing model
- Redesigned the audit filter in docker engine to read necessary parameters from docker API, convert it to CADF format, serialize it to JSON and push onto configured datastore (MongoDB, MySQL)
- Extended docker auditing to detect intrusions and threats at packet, memory, and container level

- **Indian Institute of Technology, Bombay**

May 2015 - Jul 2015

*Summer Intern*

*An e-commerce platform to empower agricultural market in rural India (supervisor: Prof. Ganesh Ramakrishnan)*

- Developed a B2C, cross-platform 'Lokacart' client app that enables farmers to connect with consumers directly
- Built analytics dashboards to track and monitor user interactions on both client and admin applications

## RELEVANT COURSEWORK

---

- **Systems:** Operating Systems, Database Management Systems, Computer Networks, Microprocessors, Computer Organization and Architecture, Unix System Programming, Compiler Design, Systems Modeling and Simulation
- **Algorithms:** Analysis and Design of Algorithms, Data Structures, Advanced Algorithms
- **ML:** Applied Machine Learning, Natural Language Processing

## SKILLS

---

- **Languages:** Python, Scala, R programming, Golang
- **Datastores:** Redis, PostgreSQL, DynamoDB, Kafka (as a data-store)
- **Frameworks:** Spark, Tensorflow, Airflow, Apache Flink

## OPEN SOURCE PROJECTS

---

- **Variational Recurrent Autoencoder:** Unsupervised, feature-based time-series clustering algorithm in pytorch
- **Troop:** A simple library to perform chunkwise processing on data.frame across multiple cores of a single machine using SNOW clusters with a low memory footprint

## AWARDS AND ACCOMPLISHMENTS

---

- Twice Most Valuable Professional (MVP) Award at Swiggy, 2020
- The Red Shift Award - Fastest learner at Noodle.ai, 2018
- Outstanding Intern of the Year Award at Noodle.ai, 2017
- Distinction Award at PESIT, semester-wise cash prizes for excellent academic performance
- **3<sup>rd</sup> place**, Citi Mobile Challenge - APAC, 2015
- Chief Minister's Scholarship, **Top 1%** candidates in Board of Intermediate Education (12th grade), 2013