

# Harsha

<https://asystemsguy.github.io>

Email : devkhv129@gmail.com

Mobile : +1-604-679-7845

## EDUCATION

---

- **University of British Columbia** Vancouver, BC, Canada  
*Master of Applied Science in Computer Engineering; Percentage: 88.5*  
*Sept. 2017 – Present*
- **Manipal Institute of Technology** Manipal, India  
*Bachelor of Technology in Computer Science; GPA: 8.86/10.0*  
*June. 2011 – July. 2015*

## EXPERIENCE

---

- **University of British Columbia** Vancouver, BC, Canada  
*Research and Teaching Assistant*  
*Sept 2017 - Present*
  - **Research Assistant - Distributed cloud-native architectures:** Research on machine learning for portfolio hedging and replication algorithms. Modeling low-risk & continuous-return strategies. Developed the python library QSTK.
  - **Teaching Assistant - Computer networks and Engineering design studio:** The online course on Coursera, had more than 100,000 students enrolled. It was featured on the 11 Alive News and the Atlanta Journal Constitution. Involved in creating assignment, exams and conducting recitation sessions. Also taught the on-campus version of the course.
- **Siemens** Bangalore, India  
*Systems Software Engineer*  
*July 2015 - July 2017*
  - **Tensorflow:** TensorFlow is an open source software library for numerical computation using data flow graphs; primarily used for training deep learning models.
  - **Apache Beam:** Apache Beam is a unified model for defining both batch and streaming data-parallel processing pipelines, as well as a set of language-specific SDKs for constructing pipelines and runners.
- **SanDisk** Bangalore, India  
*Graduate Intern*  
*Jan 2015 - June 2015*
  - **Notifications:** Service for sending email, push and in-app notifications. Involved in features such as delivery time optimization, tracking, queuing and A/B testing. Built an internal app to run batch campaigns for marketing etc.
  - **Nostos:** Bulk data processing and injection service from Hadoop to Cassandra and provides a thin REST layer on top for serving offline computed data online.
  - **Workflows:** Dataduct an open source workflow framework to create and manage data pipelines leveraging reusable patterns to expedite developer productivity.
  - **Data Collection:** Designed the internal survey and crowd sourcing platform which allowed for creating various tasks for crowd sourcing or embedding surveys across the Coursera platform.
  - **Dev Environment:** Analytics environment based on docker and AWS, standardized the python and R dependencies. Wrote the core libraries that are shared by all data scientists.
  - **Data Warehousing:** Setup, schema design and management of Amazon Redshift. Built an internal app for access to the data using a web interface. Dataduct integration for daily ETL injection into Redshift.
  - **Recommendations:** Core service for all recommendation systems at Coursera, currently used on the homepage and throughout the content discovery process. Worked on both offline training and online serving.
  - **Content Discovery:** Improved content discovery by building a new onboarding experience on coursera. Using this to personalize the search and browse experience. Also worked on ranking and indexing improvements.
  - **Course Dashboards:** Instructor dashboards and learner surveying tools, which helped instructors run their class better by providing data on Assignments and Learner Activity.

## PROJECTS

---

- **QuantSoftware Toolkit:** Open source python library for financial data analysis and machine learning for finance.
- **Github Visualization:** Data Visualization of Git Log data using D3 to analyze project trends over time.
- **Recommendation System:** Music and Movie recommender systems using collaborative filtering on public datasets.
- **Mac Setup:** Book that gives step by step instructions on setting up developer environment on Mac OS.

## SKILLS

---

- **Programming Languages:** C++, Go, Python
- **Technologies:** Cloud(AWS,Google cloud, IBM cloud), Kubernetes, Docker, Kafka, GCE