

# VIRAJ RATHOD

## DATA ENGINEER

Motivated, self-driven, and responsible Data enthusiast with 2+ years of experience in building efficient data pipelines and applications using ETL and scripting tools.

viraj.rathod@utah.edu ✉

(623)396-9127 📞

623 S 600 E, Salt Lake City, UT, 84102 📍

linkedin.com/in/virajrathod in

github.com/virajrathod 🐙

## WORK EXPERIENCE

### Data Science Intern

University of Utah, Business Intelligence Services

09/2019 – Present

Salt Lake City, UT

#### Achievements/Tasks

- Devised a Data Lineage Model consisting of over 10K tables using Scala and Spark's GraphX and visualized it using Neo4j on the centralized data server.
- Implemented multi-user security policies for Dataiku users, using Sentry and Active Directory.
- Developed a student graduation predictive model using student login data, student surveys and financial status of the student using Dataiku, Tableau and Cloudera cluster with a predictive accuracy of 81%.

### Data Engineer

Reliance Jio Infocomm Limited

07/2018 – 07/2019

Mumbai, India

#### Achievements/Tasks

- Migrated over 1M records using protobuf message format through Apache Kafka into Apache Ignite Database, automated the process using REST API and cronjobs in MS Azure saving 47% of the query time.
- Developed, maintained, and supported Continuous Integration (CI-CD) framework for code deployment based on Gitlab, fast tracking the deployment time by 90%.
- Provisioned Servers and deployed features using Ansible, ELK Stack, and maintained using Grafana, Kibana dashboards, and TICK stack for alerting mechanism.
- Actively managed and monitored cloud infrastructure on Microsoft Azure, including backups, patches, and scaling.
- Helped several project teams across various functional domains like development/testing, deployment, systems/infrastructure to ensure continuous operability to build and test systems.

### Data Engineer Trainee

Reliance Jio Infocomm Limited

01/2018 – 07/2018

Mumbai, India

#### Achievements/Tasks

- Unified over 350M customer data records from different data sources such as RDBMS, flat files for training and modeling using Apache Spark and Python.
- Automated the process of analyzing feedback, implemented Naïve Bayes classifier, and SVM by scripting CI-CD pipelines using Python and bash.
- System for OCR for IDs: Worked on OpenCV to identify text from IDs for onboarding of new employees.

## SKILLS

Python SQL C++ HDFS Spark Hadoop

Hive Tableau HTML D3.js Dataiku

Oracle 12+ Influx DB GitLab Jenkins

Apache Kafka Microsoft Azure AWS Grafana

Jupyter Nb BigQuery Matillion DOMO

## EDUCATION

Master of Science in Information Systems - University of Utah (08/2019 – Present)

Courses: Data Visualization, Business Decision Algorithms, Data Science & Big Data, Data Structures and Statistics & Predictive Analytics.

Bachelor of Science in Computer Engineering - University of Pune, India (07/2013 – 06/2017)

Courses: Data Structures, Database Management, Software Engineering, Data Mining.

## ACADEMIC PROJECTS

The Front Climbing Club (Capstone Project) – Data Warehousing (01/2020 – Present)

- Analyze customer data to develop different KPIs for the profitable sectors (gym, climbing, yoga) and visualize it for the business owners.

Earthquake Damage Visualization (Final Project) (08/2019 – 12/2019) 📄

- Effectively visualized the damage incurred to 20K buildings in San Francisco due to a hypothetical earthquake.

Tag Based Image Search (01/2017 – 06/2017) 📄

- Created flat-files containing tags from big data sets and ranked them according to the image properties and return the searched image using the tag score.

All-shops online (Massachusetts Institute of Technology - Kumbhathon) (08/2015 – 09/2015)

- Customer RFM analysis using crowdsourced data from small shops and offline stores in and around the city of Nashik.
- Developed a Web app and an android functionality for the application showing the location of every shop in Nashik.

24-Hour Analytics Case Competition (11/2019 – 11/2019)

- Utilized Domo to identify variables that correlate with safety incidents across all 2018-2019 Layton Construction work sites in order to provide KPIs and actionable insights for the management team, completed in under 24 hours.