Rohan Sadale

326 6th Ave SE, Apt 101,
Minneapolis, MN 55414
+1(612)404 4593

Email: sadal005@umn.edu
LinkedIn: https://www.linkedin.com/in/rohansadale
Github: https://github.com/rohansadale

Education

University of Minnesota, Minneapolis, MN, USA — September 2015 - Present
Master of Science, Data Science GPA: 3.75/4.00 — *(expected graduation - May 2017)*

College of Engineering (COEP), Pune, MH, India — July 2008 - July 2012
Bachelor of Technology, Computer Engineering GPA: 7.99/10.00

Technical Skills

Languages: Python, C/C++, Javascript, Shell Scripting, PL/SQL
Databases: MongoDB, MySQL, Oracle 11g
Big Data/Statistical Data Analysis: R, Tableau, Hadoop, Pig, Hive, Spark

Industry Experience

- **Veritas Technologies LLC**, Mountain View, CA, USA
  Data Analytics Intern — June 2016 - September 2016
  - Large Scale Feature Extraction, Data Cleaning and Exploratory Data Analysis of Hardware Appliances
  - Design and re-structure historical structured/unstructured datasets to make them compatible for analytical operations.
  - Build models for forecasting disk usage across appliances and predicting appliance failures.

- **Barclays**, Pune, MH, India
  Technology Analyst — July 2012 July 2015
  - Designed a streamlined Database-ETL system for generation of financial reports for APAC regulatory bodies
  - Defined and implemented underlying Database architecture, data models and methods for data extraction from heterogeneous trading databases
  - Developed data transformation and statistical analysis packages using Java and Oracle PL/SQL
  - Developed, maintained and enhanced batch framework using Perl and Shell scripting
  - Delivered Functional specifications, Scope documents, Interface documents and data mapping documents
  - Implemented and practiced agile methodology for project deliveries. Configured JIRA for the implementation

Course Projects

- **Web Scraping & Data Analysis:** Developed a web scrapper to extract restaurants' data across US. The scrapper used a novel technique to locate regions precisely using Google Maps and extracted restaurant info from those regions through Yelp. Later, performed Exploratory data analysis to visualize trends in the data and Sentiment analysis on the user reviews to suggest quality improvements for restaurant owners.
- **A Scalable Movie Recommendation System:** Built a movie recommendation system for suggesting movies to users. The project involved multiple steps such as extracting data from RDBMS, Web Logs and APIs, cleaning data using Hadoop MapReduce jobs, data filtering and transformation using Apache Hive and building scalable item-based recommender using Apache Mahout.
- **Clustering and Classification of emails:** Implemented *(from scratch) highly efficient* Ridge Regression classifier, kNN classifier, Centroid-based classifier and Spherical k-means for classifying and clustering emails into categories.
- **MapReduce like Compute Framework:** Developed a MapReduce-like compute framework for performing operations in distributed fashion. The framework would receive jobs from a client, split job into multiple tasks, and assign tasks to compute nodes which performed the computations. Proactive fault tolerance, fault detection and recovery were implemented to guarantee correct results even in the presence of node failures.
- **Distributed File System:** Implemented a Distributed File System in which clients shared files with each other. Multiple clients could read/write files at the same time and the files would be replicated (distributed) across several file servers. Achieved consistency across these files using Quorum based techniques.