

ROHAN MANOJ THAKKAR

rohan27@uw.edu

Website: 27rohan.github.io/Portfolio

Santa Clara, CA - 95050

github.com/27rohan

(206)739-4797

linkedin.com/in/rohanmthakkar

EDUCATION

UNIVERSITY OF WASHINGTON (UW), Information School, Seattle, WA *Sep 2015 - Jun 2017*
Master of Science in Information Management with specializations in Data Science & Business Intelligence (GPA - 3.7/4)

UNIVERSITY OF MUMBAI, K J Somaiya College of Engineering, Mumbai, India *Aug 2011 - Jun 2015*
Bachelor of Engineering in Computer Engineering (*First Class Honors*)

SKILLS

- **Languages:** Python (pandas, numpy, scikit-learn), Java, SQL, R, Bash (UNIX) scripting, D3.js, HTML, CSS
- **Data tools:** [MongoDB, Titan, S3, Kafka, Firebase, Cassandra] [SQL Server, MySQL, Oracle] [Tableau, Visio]
- **Frameworks and IDEs:** [AWS, Jupyter notebook (IPython), Spark (PySpark), R Studio] [XCode, Android Studio]

EXPERIENCE

VMWARE, Palo Alto, CA *Member of Technical Staff (Cloud Platform Business Unit): Jun 2017 - Present*

- Machine Learning: Feature engineering & selection, running & evaluating multiple predictive models (*Python*)
- Data Engineering: Defining schema of new tables to enable dynamic updates by scheduling cron jobs for above
- Data visualization: Creating interactive Tableau web dashboards with KPIs and actions (filters & embedded URLs)

UW MEDICINE, Seattle, WA *Big Data Developer: Jan - May 2017*

- Meshing large SQL health data from remote Linux servers & performing required analysis (*ssh, Python - pandas*)
- Mapping addresses to GPS using Google Maps API & creating interactive visualizations (*D3.js, HTML, CSS*)

CYANOGEN, Seattle, WA *Software Engineering Intern (Data Cloud Services): Jun - Dec 2016*

- Deployed the first prototype (*Python*) of App Recommender and contributed to NewsCard ranking such that they -
 - Integrate with existing AWS-based NoSQL data pipelines (Kafka and S3) to scrape historical user data
 - Perform Machine Learning (ML) using scikit-learn and Spark to estimate user interest for personalization
 - Stores top results on an AWS-based Titan graph-DB that supports querying for user through Android app
- Solely programmed and pitched a full-stack software to be integrated with the automation server such that it -
 - Performs data scraping using server APIs, logs-parsing and computations for a MongoDB server (*Python*)
 - Provides on-demand interactive online dashboard with multiple filters (*D3.js, HTML, CSS*)

LARSEN & TOUBRO TECHNOLOGY SERVICES, Navi Mumbai, India *Development Intern: Jun - Aug 2014*

- Optimized the JavaScript code - improving its speed by almost 45% and redesigned the GUI for user-friendliness

RAJ INFORMATION SYSTEMS, Mumbai, India *Data Management Intern: Aug - Nov 2013, Dec 2012 - Mar 2013*

- Delivered effective SQL querying to manage remote databases and analyzed sales figures to increase profit by 5%

RECENT PROJECTS

TICKET GENIE (Capstone with iLink Systems, Redmond, WA) *May 2017*

- Coded the machine learning modeling behind all the predictions and persisted it for REST API requests (*Python*)

FREEFOOD - <https://github.com/27rohan/Free-Food> *3 - 13 Dec 2016*

- Coded in Swift 3 the real-time JSON data flow (read and write) between the iOS app and Firebase server

LEARNING FROM TWITTER (DubHacks 2016) - https://github.com/27rohan/Twitter_Learn *15 - 16 Oct 2016*

- Visual recommender system that for a valid Twitter user, suggests top 10 similar users and words to tweet about
- Technologies: Twitter API, Microsoft Cognitive API, Firebase, Python, D3.js, HTML, CSS, AWS

CHASE IT - https://github.com/27rohan/Chase_it *Apr - May 2016*

- Built an addictive native Android game with Android Studio using Java on back-end and XML on front-end
- Features: Multiplayer mode (Wi-Fi direct), touch screen and motion sensor recognition, local leaderboard

PET PROJECTS IN DATA SCIENCE - <http://27rohan.github.io> *Dec 2015 - May 2016*

- Understanding impact of words in critics by reviewers on the ratings using supervised ML on Rotten Tomatoes data
- Achieved store-sales forecasting (*ranked in top 10%*) for 45 Walmart stores as part of a Kaggle competition
- Implemented top ML techniques in Python scikit-learn on 5 other problems, each with a different prediction focus
- Performed web scraping using BeautifulSoup on the course catalog data of IIT in CSV & pushed as SQL to AWS