ROHAN MANOJ THAKKAR

rohan.m.thakkar@gmail.com   Mountain View, CA - 94041   (206)739-4797

LinkedIn: linkedin.com/in/rohanmthakkar   **Website:** 27rohan.github.io/Portfolio   **Github:** github.com/27rohan

EDUCATION

M.S., University of Washington (UW), Seattle, WA   *Information Management: Sep '15 - Jun' 17*

- **Data Science**: Foundational Statistics, Machine Learning (ML) & Data Scaling
- **Business Intelligence:** Data Warehousing, Analytics & Reporting
- **Native Mobile Applications Development:** Android app development, iOS app development

B. Eng., University of Mumbai, India   *Computer Engineering: Aug '11 - May '15*

SKILLS

- **Languages:** Python, SQL, Java, R, Bash (UNIX), D3.js, XML, HTML, CSS
- **Tools:** [MongoDB, Titan, S3, Kafka, Firebase, Cassandra] [Oracle, SQL Server, MySQL] [Tableau, Visio]
- **Frameworks and IDEs:** [AWS, Jupyter notebook, PySpark, R Studio] [XCode, Android Studio]

EXPERIENCE

VMWARE, Palo Alto, CA   *Data Science Engineer (FT): Jun '17 - Present*

- **Data Science & Analytics:**
  - Experimenting with ways to identify flaky tests across the vSphere suite *(Python, SQL, Tableau)*
  - Built a ML system to predict resolution time of vSphere bugs *(Python, sklearn, Keras, SQL)*
  - Devised an allocation system for distributing key resources across teams under business unit *(Python)*
  - Published multiple interactive reports to portray KPIs of productivity to management *(Tableau, SQL)*
  - Redesigned the ETL of most of the team's reporting and data warehousing for enhanced processing *(SQL)*
  - Analyzed support-requests data w.r.t. bugs to identify ways of lowering resolution time *(Python, SQL)*
- **Software Engineering*:***
  - Adding new features and resolved key issues in the CI/CD Jenkins pipeline for 2 concurrent projects:
    1. VMware Cloud on AWS   2. vSphere new delivery model

UW MEDICINE, Seattle, WA   *Big Data Programmer (PT): Jan - May '17*

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

CYANOGEN, Seattle, WA   *Software Engineering Intern, Data Cloud Services (FT): Jun - Dec '16*

- **Data Cloud team:** Co-developed 2 cloud-based ML services (App Recommender, NewsCard ranking) that -
  - Perform machine learning to personalize ads and articles based on usage data *(sklearn, PySpark)*
  - Integrate with AWS-based NoSQL data pipeline to scrape historical user data *(Kafka, S3, TitanDB)*
- **Automation team:** Developed a full-stack software that integrated with Lava automation server and -
  - Scrapes log data using server APIs, parses for computations that go to on a MongoDB server *(Python)*
  - Provides an on-demand online interactive dashboard with multiple filters *(D3.js, HTML, CSS)*

RAJ INFORMATION SYSTEMS, Mumbai, India   *Data & Software Engineer (FT): May - Aug '15*
          Data Management Intern (PT): Aug - Nov '13, Dec '12 - Mar '13

- Created ETL packages for a SQL DWH using MS-BI Suite & analyzed sales figures to increase profit by 5%

KEY PROJECTS

TICKET GENIE (Capstone with iLink Systems, Redmond, WA)   *May '17*

- Coded a ML-based classifier for real-time categorization of support tickets *(HTTP, CSS, Python: sklearn, Flask)*

FREEFOOD - https://github.com/27rohan/Free-Food   *3 - 13 Dec '16*

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

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

- AWS-based visual recommender system that for a valid Twitter user, suggests top 10 similar users and tags

PET PROJECTS IN DATA SCIENCE - http://27rohan.github.io   *Dec '15 - May '16*

- Achieved store-sales forecasting *(ranked in top 10%)* for 45 Walmart stores as part of a Kaggle competition
- Performed web scraping using Beautiful Soup on the catalog data of IIT in CSV & pushed as SQL to AWS