ROHAN MANOJ THAKKAR

rohan.m.thakkar@gmail.com   Tacoma, WA - 98402   (206)739-4797
linkedin.com/in/rohanmthakkar   27rohan.github.io/Portfolio   github.com/27rohan

EXPERIENCE

ZILLOW GROUP, Seattle, WA   *Machine Learning Engineer, User Understanding AI (Jan '20 - Present)*

- Created a new **decision engine** to automate multivariate experimentation & optimization across ZG
  - Already onboarded 2 experiments for a collective annual revenue of $15M+ in marketing campaigns
- Collaborated with scientists, engineers & PM to productionize ML models at 100M+ user scale
  - Time to launch new ML models down from ~2 weeks to ~2 days
  - Time to retrain new ML models down from ~1 weeks to ~1 day
  - Time to perform feature selection & hyperparameter tuning down from ~1 week to ~1 day
  - Time to launch automated scoring & measurement of ML models down from ~1 week to 1 hour
- Created an ingestion service to export all user scores & experimentation decisions to 4 ZG platforms daily
  - ZG datalake, ZG User Profile, Tealium & Krux (Salesforce Audience Studio)
- Team's **Airflow champion** dedicated to improving performance & deprecating legacy Airflow DAGs
- Launched & maintained AWS **infra** for a team of 15+ to prototype & productionize big data workloads
  - Sustainable process for creating & using custom AMIs to save ~12% in AWS costs (~50k/year)
  - Automated backup & restore to prevent deletion of work saved on EMR during restarts
  - Jupyter notebook launching script with Pypspark preinstalled & users tagged on resource manager
  - AWS roles with necessary datalake permissions, cluster monitor, Git support & more
- Joined ZG Data Infra team as a **Terraform admin** to continually review & apply 100+ ZG-wide MRs
- Added features for **Facebook Ads** Insights ETL & cut down FB dynamic costs by 30% via dynamic dedup
- Kickstarted & led team's **oncall** processes, including weekly handoff meetings & grooming OPS tickets
- Led ~50% of team's **Sprint** planning & **Retro** w/ stakeholder communication for key timelines & risks
- Mentored 1 MLE & 1 intern (now FTE scientist) – with plans to mentor another 2 FTEs over next ~6 months
  Technologies: Airflow, Python, Spark, AWS, Hive, Terraform, Jupyter, Facebook Business

VMWARE, Palo Alto, CA   *Member of Technical Staff (Jun '17 - Jan '20)*

- Developed multiple dashboards with automated data pipelines from different sources to present some KPIs
- Led several efforts in data science & analytics to help enhance visibility into engineering productivity
- Built features for continuous integration and delivery of core projects under the vSphere stack
  Technologies: Python, Bash, SQL, NoSQL, YAML, Tableau

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

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

CYANOGEN, Seattle, WA   *Software Engineering Intern (Jun - Dec '16)*

- Co-developed 2 cloud-based ML services (App Recommender & NewsCard ranking)
- Also created the end-to-end data pipeline from data extraction to knowledge graph publishing for the above
  Technologies: AWS (Kafka, S3, TitanDB), Python, Bash, D3.js, HTML, CSS

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)*