

ROHIT LOKWANI

☎ 206-945-4035 ✉ rlokواني@uw.edu 🔗 [linkedin.com/in/rohitlokواني/](https://www.linkedin.com/in/rohitlokواني/) 🌐 github.com/rohitl17

Education

University of Washington

Masters of Science in Data Science - GPA: 4.0/4

March 2023

Seattle, WA

Pune Institute of Computer Technology

Bachelor of Engineering in Computer Engineering - GPA: 9.2/10

June 2018

Pune, India

Experience

Seagate Technology

Data Science, Machine Learning and Software Intern

Bloomington, MN

Jun 2022 – Sept 2022

- Working on self-supervised and supervised techniques for detecting defects in sliders for hard disks. Performed hypothesis testing (Welch T-test) and correlation experiments to derive prediction confidence metrics for continual learning. Initial production deployment shows reduced human effort for retraining image selection by 80%
- Building a Evidently, Prometheus and Grafana-based application for monitoring data and prediction drift using Wasserstein distance test which will handle a load of 3M images/day to detect unforeseen excursions in the slider process. Additionally contributed to 2+ projects (tools, POCs) in ML production workflow

DeepTek Inc

Data Scientist (Tech Lead)

Pune, India

Mar 2019 – Aug 2021

- Designed DS product roadmaps and timelines, leading 6 data scientists for developing pathology detection ML workflows in CT (3D) model research, annotations, optimization, engineering and internal/external evaluations following agile methodology. Part of group that pitched technical progress to VC investors.
- Managed communication with Microsoft Research and NTT Data teams for a research project on using radiological and laboratory parameters to build COVID-19 severity predictive models. Led the project development right from framing the hypothesis, performing EDA, feature engineering and correlation and tree-based model experiments.
- Independently spearheaded the model internal evaluation process and tool design and development, improving models upto 47% precision by conducting reviews, error analysis and deriving actionable insights; collaborated with 3 cross-functional teams (16 members)
- Managed 2+ recruitment drives, interviewed 60+ candidates, hired and mentored 5 joiners, published 5 papers, conducted functional and hypothesis tests for the product and research papers; organized 6+ team talks and brainstorming sessions to adopt new techniques for DS problems
- Joined as their 4th employee, instrumental in contributing to company's growth from inception to 150+ hospitals and 10M+ USD in revenue

Associate Data Scientist (Machine Learning Engineer)

March 2019 – July 2020

- Developed and integrated 11+ machine learning models in production (Sensitivity = 90%, Specificity = 80%), contributed to 8 winning POCs requiring quick and responsible experimentation including Oracle, GE Healthcare, in turn fostering business relationships with them. Consequently rated as the best performing data scientist
- Designed web-based, RESTful APIs in Python, including data ingestion and retrieval module, inference logic, and optimized the API and deep learning models for the microservices architecture, supporting 20+ AI models and 8+ microservices, making it 15x faster and 5x memory efficient, ensuring minimal downtime and prediction latency
- Designed the product CI/CD pipeline using Docker, Jenkins, Kubernetes and Spinnaker, which handles upto 1.2 M images per day
- Built a TB detection model which got us our first AI client, CHAI (Clinton Health Access Initiative), improving the sensitivity of TB detection by 35%, using augmentations and cyclic learning rates. It has screened 150,000+ patients in Chennai, finetuned versions of this model lead to 2nd rank amongst 20 companies in TB research
- Designed a NoSQL database (MongoDB) for storing CT scans, storing 15,000+ CT scans (6M+ Images)

RackWare Inc

Trainee Software Engineer

Pune, India

July 2018 - March 2018

- Improved product reliability by 30% by predicting ZFS storage using smoothed moving average; developed cloud-agnostic features contributing to data sync pipeline for cloud APIs (Azure, AWS, GCP, OCI) serializing data transfers using Google protobufs

Projects

Stripe Customer Behavior Analysis | [Report](#)

UW | Feb 2022 - Mar 2022

- Analyzed over 1 million simulated transactions. The goals of the analysis was to use transaction data, segment customers and suggest offers and benefits to the customers to keep them engaged with Stripe. Another goal was to predict the risk of merchants discontinuing software use in the future.

Analysis of Cryptocurrencies | [Link](#)

UW | Dec 2021 - Jan 2022

- Working on identifying patterns, trends and maximize returns for top 20 cryptocurrencies (95% market capitalization). The analysis aims to help a new investor devise an informed investment strategy using inferences from statistical tests like t-test and cross-correlation using R.

Uber/Lyft Dynamic Pricing Prediction | [Project](#)

UW | Jan 2022-March 2022

- Designed and developed a web application for predicting dynamic price for uber and lyft using linear regression. It uses flask and public APIs for geospatial(Google Maps), weather and traffic data. A feedback loop retrains the models timely using a CRON scheduler.

Technical Skills

Languages: C++, C, Python (PyTorch, Scikit, Tensorflow, Pandas, Keras), Bash, R, Java, MySQL

Tools and Frameworks: GCP, Azure, AWS, Spinnaker, Kubernetes, Django, Flask, Git, Jira, Jupyter, JMP

Technologies : Google Protobufs, ZFS, Apache Kafka, Docker, Spark, SystemML