Usman Khaliq

Portfolio: usmankhaliq.com | Email: usman.khaliq91@gmail.com

SKILLS

- **Programming Languages:** Python, SQL, R
- **Engineering:** Airflow, ETL pipelines, Git, Makefile, package development, Unix, python testing frameworks(pytest)
- **Data Visualization Softwares and libraries**: Looker, Redash, Tableau, Retool, ggplot2, seaborn, R Shiny
- **Other Skills**: Data Engineering, Exploratory data analysis, experimental design, statistical inference, machine learning

RELEVANT WORK EXPERIENCE

Coinbase — **Oct 2021 – June 2022**
Data Scientist — Boston, MA
- Defined core retail market share metrics and created dashboards to assist with deep dive analysis of fee and revenue drivers
- Wrote **data pipelines** in Airflow for automating product metric generation
- Launched **A/B tests** to measure the impact on **customer engagement** and **company revenue** from high visibility products (retail advanced trading, gains and losses)
- Developed **machine learning models** for revenue forecasting

Overjet — **June 2020 – April 2021**
Data Scientist — Boston, MA
- Defined **product success metrics**, created **data warehouse infrastructure** to track metrics and **generated reports** to showcase the impact of Overjet's disease diagnosis AI to dental clinics. Increased the return of investment for a major dental organization in New England by 33.6X
- Analyzed **structured and unstructured clinical data** and deep learning model predictions to define **product requirements** and launched **two new products** (chart audit and dental restorations) over a **7 month** period
- Developed a clinical **text processing service** that improved the **specificity** of the tooth decay prediction models by **80%**
- Wrangled **5 years of patient data** to **identify potential fraud** in dental procedures for a dental organization in the US
- Developed **data ingestion pipelines** in Python that **reduced time of onboarding** new clinics from **5 days** to **2 hours**

Verily Life Sciences — **June 2019 – Sept 2019**
User Experience Research Intern — Cambridge, MA
- Conducted **qualitative research** on developing an end-to-end product to diagnose and manage sleep apnea. Convinced the **product team** to **prioritize development** of an audio based guided meditation module for sleep apnea patients

Systems Utilization Research for Stanford Medicine (SURF) Lab — **April 2019 – Dec 2019**
Research Assistant — Stanford, CA
- Built a **data visualization application** using **Shiny R** that implemented different variants of the Medicare for All model developed by Schulman & Milstein. Published in Health Management Policy and Innovation Volume 4
- Developed a **clustering algorithm** from **electronic health records** and patient registry data for patients with ventricular septal defect to implement a **target-based care program** at the Stanford Children's Hospital

EDUCATION

Stanford University — **2018-2020**
MS Design Impact
Stanford Data Lab: Taught by Hadley Wickham (Chief Data Scientist at RStudio).
- Developed extensive proficiency in exploratory data analysis techniques using the tidyverse packages in R
- Conducted Network analysis on FiveThirtyEight's dataset of 3 million Russian Troll Tweets to determine how different categories of troll accounts coordinated before major political events.

Ghulam Ishaq Khan Institute of Engineering Science and Technology — **2011 - 2015**
BSc Computer Engineering (Graduated in Top 10% of Class. Dean's Honor List in 3rd and 8th Semesters)