

Omar Ali

San Francisco, CA · Omarh90@gmail.com · github.com/Omarh90

Highlighted Projects

EHR Clinical Database Harmonization: Python, SQL, AWS, NoSQL, XML Melissa Informatics
· Overhauled ETL legacy pipelines for tenfold SQL DB scaleup & extensive schema changes December 2018
· Remediated data integrity for patient IDs, creating graphics to illustrate data corruption cause and scope

NLP Histopathology Scan OCR & Context Extraction: R, SQL, ML Melissa Informatics
· Wrote algorithms for medical file classification and NLP parsing with AI & advanced statistics March 2019
· Compared multidimensional feature clusters to differentiate medical forms using matrix algebra

Image Processing Numerical Integration Application: Python McCampbell Analytical
· Wrote application to numerically integrate graphic to correct erroneous software calculations August 2018
· Includes GUI (tkinter OOP), automatically generating report using multinomial calibration and OCR

Experience

FreightVerify: November 2019 – May 2020
AI Engineer San Francisco, California
· Wrote AWS geospatial aggregation mapping tool with anomaly detection SQL queries & Python graphics
· Evaluated distributed DB infrastructure tools and Cloud-based AI, including Hadoop, AWS, & GCP

Code for San Francisco: July 2019 – October 2019
ShelterTech Dev & Analytics Team (Volunteer)
· Configured search engine for online website hosting homelessness resources in SF, including accessibility
· Critiqued CA Department of Motor Vehicles website through DMV Strike Team user interface testing

Melissa Informatics: November 2018 – June 2019
Knowledge Engineer (Data Engineer) Berkeley, California
· Integrated EHR databases, optimizing NoSQL queries, scaling pipelines x10, & maintaining legacy code
· Applied unsupervised learning to healthcare medical files to automate classification and field recognition
· Automated histopathology content parsing using R and SQL, including ML, NLP, & OCR analysis

McCampbell Analytical: August 2017 – October 2018
Research & Development Analytical Chemist Pittsburg, California
· Wrote Python image processing program to correct faulty software calculations, including user interface
· Guided quantitative development of scientific calculations for novel experimental setups, including DB

Northwestern University: September 2016 – August 2017
Quantitative Bio-Element Imaging Center, Research Technologist II Evanston, Illinois
· Prepared clinical early phase I FDA statistical validation setup for gene-therapeutic glioblastoma trial
· Trained graduate students in data analysis and troubleshooting in R, and maintained lab's WordPress site
· Applied multidimensional time series data processing for laser ablation & high res mass spectrometry

University of Illinois, Urbana-Champaign: Various Titles Held August 2008 – August 2016
Center for Chemistry and Technology, Assistant Analytical Chemist March 2014 – August 2016
· Wrote dynamic, user-friendly VBA script to parse, seek, & compute QC criteria in multi-dimensional data
· Increased throughput four-fold for ICP-OES analysis, including streamlining data processing

Education

University of Illinois at Urbana-Champaign: December 2012
Bachelor of Science in Mathematics GPA: 3.24/4.00
Bachelor of Science in Chemistry

Technical Skills

Programming: Python (pandas, pyspark, keras, tkinter), R, MATLAB, C, VBA · **DAG:** Airflow, Knime
Database: MySQL, PostgreSQL, NoSQL · **Cloud:** AWS (RDS, Sagemaker), GCP · **OS:** Linux, Windows
Stats: Machine Learning, AI/Deep Learning, Logistical Regression, Linear Algebra, Time Series, NLP