Roger Qiu

Dallas, TX | rogerqiu7@gmail.com | 832-349-0490

[LinkedIn](#) | [Portfolio](#)

EDUCATION

UNIVERSITY OF SAN DIEGO | M.S. APPLIED DATA SCIENCE | GPA: 3.8   09/2022 - 08/2024

Coursework: Statistics, Data Science Programming, Foundation of Data Science, Machine Learning, Applied Data Science for Business, Applied Data Mining, Applied Predictive Modeling, Machine Learning, Applied Time Series Analysis

SAINT LOUIS UNIVERSITY | B.S. HEALTHCARE ADMINISTRATION, INFORMATION TECHNOLOGY MANAGEMENT   08/2015 - 05/2019

Coursework: Health Operations Management, Enterprise Applications Development, Systems Analysis and Design

EXPERIENCE

COHERE HEALTH   Boston, MA

Data Engineer   02/2023 - Present

- Developed and optimized dozens of Python and Spark ETL pipeline jobs, reducing job runtime in Airflow by up to 90% by redesigning and creating more efficient code.
- Automated in-depth exploratory data analysis (EDA) and preprocessing of millions of records daily within minutes by creating dozens of packaged functions, stored in our GitHub.

NATIONAL OILWELL VARCO   Houston, TX

Financial Analyst   10/2022 - 02/2023

- Leveraged Alteryx and SQL to create dozens of automated data pipelines and workflows, saving hundreds of hours of manual ETL processes.
- Increased management decision making capabilities by developing and presenting Power BI dashboards to report financial data efficiently using visualizations and dashboard summaries.

WERNER CO.   Itasca, IL

Financial Analyst   10/2021 - 10/2022

- Modeled and presented monthly financial data from company income statements using IBM PA (Cognos TM1), Excel, and Power BI, guidings departments to millions in savings of quarterly budget

MERRIL LYNCH   Chesterfield, MO

Financial Advisor   05/2019 - 10/2021

PROJECTS

DATA SCIENCE | MACHINE LEARNING | SOFTWARE DEVELOPMENT

Breast Cancer Prediction

- Data science project using machine learning techniques to diagnose breast cancer such as EDA, data pre-processing and multiple models. Hyperparameter tuning is used to optimize these models, and feature selection techniques.

Patient Management System

- Patient management database created using Java and Java Swing for the GUI.

SKILLS

Python, Apache Spark, Apache Airflow, AWS Athena (SQL), Github, AWS Sagemaker, AWS Elastic MapReduce (EMR), Jira, Supervised Machine Learning, Statistical Modeling, Java