Muhammad Saad Khalid

(347)-822-8693 | muhammadsk123456@gmail.com | Brooklyn, NY | Github | LinkedIn

Education

Baruch, City University of New York (CUNY), Brooklyn, NY
BS in Computer Science | GPA: **3.92**/4.0 Expected December 2026
Minor in Mathematics
Relevant Coursework:
Data Structures and Algorithms, Programming, Database Management Systems, Statistics, Machine Learning

Skills

Programming Languages: Python, SQL, VBA, C++
Libraries: Pandas, NumPy, SciPy, Scikit-Learn, Matplotlib, SeaBorn, PyTorch, TensorFlow, PennyLane
Database Management Systems: Microsoft Access, MySQL, PostgreSQL
Data Visualization: Tableau

Relevant Experience

Researcher at University of Tennessee | GitHub May 2025 - August 2025
Python, PyTorch, TensorFlow, PennyLane, NumPy, SciPy

- Implemented Quantum Neural Networks (PennyLane, TensorFlow) and classical Neural Networks (PyTorch) to predict optimal power flow solutions for complex energy grid scenarios, achieving a classical model MSE of 0.001.
- Optimized quantum neural network architecture through angle embedding and output measurement design, achieving 50% reduction in computational resource requirements.
- Evaluated model scalability and computational efficiency across various power system sizes, modeling large scale city energy systems from 3 - 14 sized graphs, demonstrating robust performance for increasing grid complexity.

College Tutor August 2024 - December 2024
CUNY - Baruch College

- Simplified complex concepts in computer organization, data representation, and Python algorithm development, providing tailored 1:1 or 1:2 tutoring.
- Facilitated weekly hands-on sessions for 10-15 students over a semester, solidifying understanding of algebraic concepts and enhancing problem-solving skills.

CUNY Tech Prep Fellow July 2025 - Present

- Selected for competitive data science fellowship with students from across 11 CUNY senior colleges where fellows create technical projects using tools such as Python, Jupyter Notebook, Pandas, Numpy, Scikit-learn, and SQL.
- Participating in weekly courses and learning industry best practices for exploratory data analysis (EDA), feature engineering, data collection and processing, statistical modeling, data visualization, ML techniques, and big data.

Projects

Automated Email Prioritization System | Github
Python, Groq AI August 2025 - September 2025

- Built an AI automation workflow leveraging Large Language Models (LLM's) to summarize and classify unread emails by importance.
- Applied Natural Language Processing (NLP) and parallel processing to optimize analysis, reducing processing time by ~75%.
- Developed a scalable background data pipeline with scheduled execution, enabling real-time automated reporting and CSV output without manual intervention.

COVID Early Policy Analysis | Github; Tableau
Python, PostgreSQL, Tableau June 2025 - July 2025

- Engineered ETL pipelines to extract, transform, and load 85,000-row global COVID-19 dataset (2020-2021), converting daily records into standardized weekly time-series with integrated epidemiological and demographic variables.
- Identified 10 most similar countries to the US through Euclidean distance, conducted three-phase time-series analysis showing the US ranked 8th-10th among peers with outcome scores 11-41x worse than top performers.
- Designed interactive dashboards to visualize early COVID-19 case and death trends across countries, including time-series progression for the US and demographically similar nations.

Data-Driven Health Risk Analysis
Python, Scikit learn February 2025 - May 2025

- Developed a Decision Tree Classifier achieving 90% recall for early heart disease prediction.
- Applied Linear Regression to identify strong positive correlations between BMI and heart disease.
- Applied K-Means and Hierarchical Clustering to segment patient populations, revealing distinct at-risk groups and validating BMI-heart disease correlations for targeted intervention strategies.