

Qualification Summary

Proactive and results-driven professional with a blend of Statistics & Applied Mathematics and expertise in maintaining and expanding existing data collection and delivery platforms.

Skilled in conducting in-depth analysis and identifying requirements to enhance user experience. Highly motivated to develop and implement analytics applications to transform data into meaningful information. Notable acumen in analyzing data and interpreting patterns and bringing machine learning solutions to AI products with keen understanding of Statistics, Mathematics and data science and software engineering practices. Ability to support delivery of high-profile projects and apply cutting-edge techniques to find insights as well as define solutions for data-driven decision making in a cost effective and scalable manner. Well-prepared to work with IT teams in exploring data, building/implementing predictive models, and performing data visualizing and reporting to strategically support business. Equipped with strong communication, multitasking, and learning skills.

Data Management & Acquisition | Machine Learning | Data Science Engineering | Project Management | Statistical Efficiency | Data Visualization/Mining | Database Management | Teamwork & Leadership | Strategic Planning & Execution

Technical Proficiencies

Technologies: TensorFlow, Scikit-Learn, SciPy, Optuna, OpenCV, NLTK, PyMC3, Tableau, React, Express, Linux, Asana, Google Cloud Platform (Vision, Tables, BigQuery, Vertex AI, Cloud Storage), Amazon Web Services (SageMaker, Lambda, EC2, Amplify, RDS, S3), Palantir Foundry, Firebase, DbVisualizer, Jira, Confluence

Languages: Python (Jupyter Notebook, Google Colaboratory), SQL (MySQL, PostgreSQL, Vertica), JavaScript, C++

Experience

Fiserv - Berkeley Heights, NJ

2023 – Present

Data Scientist

Led the development of an XGBoost model used to prioritize and automate the decisioning of fraud and credit alerts. Significantly reduced analyst alert queue volume (more than \$100,000 worth of work annually). Supported the migration from legacy systems (FICO Blaze Advisor) to modern machine learning environments (AWS). Created and maintained industry based statistics on merchant processing. Crafted new rule based alerts to increase alert coverage. Augmented the standing merchant transaction limit assigning logic to decrease alert volume and increase bad capture rate

WebMD – New York, NY

2022

Data Science Intern

Provided insights for clients by determining best doctors for drug advertisement through exploratory data analysis, linear regression, and k-nearest neighbors regression. Developed lookalike audiences using k-means clustering and elbow curve method to find optimal number of clusters and visualized the results in Tableau in order to clearly explain findings to team members. Adhered to HIPAA privacy standards while handling data.

- Devised a medical doctor and nurse verification system by utilizing Google Cloud Platform and more than 30,000 data points provided by data scraping contractors.

Convertli – Hoboken, NJ

2021 – 2022

Software Engineer

Implemented meticulous approach in creating and maintaining product sales and views NoSQL Firebase databases. Allowed users to track conversions and sales by designing merchant facing data analytics panels through best engineering practices.

- Succeeded in increasing and tracking customer conversion rate by producing Shopify apps.
- Developed database schemas to illustrate data flow, usage, and acquisition in a clear manner.

Education & Credentials

Master of Arts in Statistics & Applied Mathematics, 2023 | BA/MA - Major GPA: 3.75

City University of New York, Hunter College – New York, NY

Acquired basic and advanced statistical knowledge with focus on areas of data science and decision making. Built advanced proficiency in computer languages and skills to leverage mathematical theories and principles to solve real-world challenges.

- Learned fundamentals of Data Science & Machine Learning, Adv. Machine Learning/Deep Learning/NLP, Numerical Methods, Stochastic Optimization, Modeling & Visualization, Adv. Probability Theory, and Measure Theory.

Bachelor of Arts in Mathematics, Minor in Computer Science

City University of New York, Hunter College – New York, NY

- Linear Algebra, Real Analysis, Ordinary Differential Equations, Calculus 1-3, Topology, Abstract Algebra, Software Analysis, SQL Programming, Discrete Probability, Applied Statistics

President | South Asian Culture Club

- Organized and hosted in-person events for more than 350 attendees
- Created Tableau dashboards to explain event finances and insights to executive board members
- Recruited more than 200 new members and controlled and allocated more than \$10,000 worth of funds

Key Projects

Face Recognition with Siamese Neural Networks

- Minimized cost from contrastive loss using stochastic gradient descent
- Optimized for precision to reduce false positives and achieved 85% precision on testing data
- Utilized a Siamese Neural Network to create embedding vectors in representation space for 2 images at a time
- Implemented a one-shot learning approach to create a model trained off more than 10,000 pairs of images to determine persons' face through pair of images

An Analysis of Pakistani Homes

- Analyzed the Pakistani housing market using exploratory data analysis and machine learning to predict home prices
- Built and compared models using linear regression, polynomial regression, k-nearest neighbor's regression, decision tree regression, and random forest regression and ranked models as per coefficient of determination

Credit Card Fraud Classification

- Investigated credit card transactions and used machine learning methods to identify fraud transactions
- Used sampling techniques such as undersampling to balance the dataset for best results
- Trained models such as decision trees, random forest, and support vector machines and optimized using cross-validation and grid search to fine tune hyperparameters

Gasoline Demand Time Series Forecasting

- Explored patterns, trends, and seasonality of US gasoline demand data
- Employed ARIMA, XGBoost, Facebook Prophet, and Recurrent Neural Network to forecast future gasoline demand and compared models using mean squared error

Planetarium Arcadium

- Led a team of 4 people to create a full (PERN) stack orrery web app
- Stored dataset in a PostgreSQL database hosted in AWS EC2 instance served through Express server
- Built web app using JavaScript, React, and Node to enable user to search/visualize distant stars and orbiting planets