Mitchell Krieger

GitHub | Blog | LinkedIn
New York, NY | (248) 497-1993 | mitkrieger@gmail.com

TECHNICAL SKILLS

- **Python:** pytorch | pyspark | scikit-learn |numpy | pandas | matplotlib | plotly | dash | scipy | statsmodels
- **Machine Learning & AI:** classification | regression | deep learning | decision trees | clustering | entity resolution
- **Tech Stack:** SQL | C++ | Databricks | Vertex AI, BigQuery & Google Cloud Platform | Tableau | dbt | Git | Excel

RELEVANT EXPERIENCE

Data Scientist, Middesk, New York, NY — September 2022 - August 2024

- Developed, trained, and managed machine learning models to enhance product offerings and decision-making.
  - Iteratively trained and evaluated ML models for Middesk's entity resolution engine, achieving a 0.95 F1-score in resolving 102M+ profiles into business identities (clusters of data attributes, documents, and records).
  - Increased the accuracy of Middesk's industry classification product by 5%, a hierarchical set of ML models that classify businesses into 300+ unique industries, while also implementing cost-saving measures leading to a 30% reduction in cost.
  - Applied academic literature to guide ML development, implementing research-based algorithms tailored to business needs.
  - Improved Middesk's matching model for finding businesses based on customer submitted data, reducing manual reviews by 2% while continuing to return high-quality results to customers.
- Partnered with leadership and key stakeholders to define and track metrics that support strategic decisions.
- Enhanced data quality, accessibility, and metric calculation consistency by creating a library of 30+ analytics tables

Data Scientist / Business Intelligence Analyst, BlockFi, New York, NY — May 2021 - August 2022

- Delivered data-driven solutions to optimize revenue and mitigate risk in cryptocurrency markets using statistical analysis and machine learning.
  - Assessed the impact of trading spread changes on 1 million retail clients using statistical testing and recommended increasing spreads, resulting in a 7% increase in retail trading profits.
  - Developed an interactive tool using Plotly-Dash & Databricks for sales reps to manage 3,000+ high-net-worth clients, driving $50M+ in additional digital asset inflows.
  - Built BlockFi's analytics infrastructure using Databricks and Spark, creating bronze/silver/gold layer Delta Lakehouse tables.
  - Designed and maintained dashboards in Tableau, Databricks, and Plotly Dash to track KPIs and communicate insights.
- Collaborated with stakeholders across the finance, AML & compliance, fraud, and client service verticals to understand data needs and business use cases to create various data-driven solutions
  - Presented at Chainalysis LINKS conference on utilizing on-chain data across different use cases (**Watch on YouTube**)

SELECTED ADDITIONAL EXPERIENCE

Treasurer, Director, and Curricular Consultant, L!FE Leaders, Detroit, MI/Remote — January 2018 - Present

- Implement curricular & institutional structures, helping to establish L!FE as a growing non-profit start-up
- Handle L!FE's budgets and financial models, while recruiting, training & managing staff

Academic Advisor, New York University, New York, NY — October 2018 - August 2020

- Created and implemented an overhaul of the advising program serving 1400 students per year
- Advised a caseload of 250+ undergraduates, providing guidance on academics, career development, and student life

Mathematics Faculty, CITYterm at The Masters School, Dobbs Ferry, NY — August 2014 - June 2017

- Taught 11th & 12th grade math ranging from Algebra II to AP Calculus BC to ~180 students
- Created a dynamical systems project where students created original mathematical models using Python and Euler's Method
- Co-taught and developed the Urban Core curriculum (Urban Design, Literature & History of NYC) using experiential learning
- Served in loco parentis to 30+ boarding students per semester providing care and fostering a positive environment

TECHNICAL PROJECTS

Rebalancing the Citibike System using Time Series Analysis - Github | Walkthrough | Dashboard Demo
Using time series analysis & clustering to identify how to distribute bikes across the system to best meet rider demand

- Mined data and analyzed trends in the 17m+ Citibike trips across 1,900+ stations made in 2018
- Built forecasting models using Facebook Prophet to predict the number of bikes at a station over time. Then extracted seasonality from models & employed clustering techniques to classify stations as pooling, draining, or balanced stations

EDUCATION

Cornell University (Cornell Tech), New York, NY — May 2025
Master of Engineering | Computer Science | Merit Scholar

Hampshire College, Amherst, MA, — May 2014
Bachelor of Arts | Educational Technology and Curriculum Development, Mathematics Education