

Luke Sullivan --- Jr. Data & Machine Learning Engineer / Analyst / Scientist
theeluke13@gmail.com, +1 (484)-510-0194, New York, NY

[LinkedIn](#), [GitHub](#)

SUMMARY

Driven Data Scientist with 2+ years of experience in data analysis and machine learning environments. Placed on a fast-tracked career path, completing a bachelor's degree in three years and now advancing with a Master's in Data Science at NYU. Combines a strong understanding in Python and SQL with hands-on experience in predictive modeling, data visualization, and agile project management.

EXPERIENCE

- 1. Data Scientist @ CSL Plasma - King of Prussia, PA** MAY 2024 - AUG 2024
 - Developed, tested, and validated ARIMA models to **forecast** daily plasma donations for **17%** better optimized resource allocation.
 - Enhanced predictive model accuracy by cleaning and analyzing **time-series data** from 10 European plasma centers. Engineered **key features** including seasonality, weekly trends, and promotional timings.
 - Worked in a **cross-functional team** of Data engineers, analysts, scientists and managers.
- 2. Data Analyst @ Mindcracker - Philadelphia, PA** DEC 2023 - MAR 2024
 - Presented a data-driven analysis dashboard (Power BI) to shareholders that identified high-growth industries for the talent acquisition department to target, leading to a **25%** increase in qualified candidate leads in the following quarter.
 - Led development of a multi-page website using **HTML, CSS, & JavaScript** utilizing **Scrum** agile methodologies, finishing web development project **on time**, following proposed sprint schedule.
- 3. IT Support Technician @ Edgewood University - Madison, WI** OCT 2022 - MAY 2024
 - Provided IT Support to the entire university, resolving and explaining over **400** technical tickets to non-technical individuals.

EDUCATION

- 1. Master's in Data Science @ New York University** SEPT 2024 - MAY 2026
 - **GPA:** 3.4 (Expected)
 - **Courses:** Database Systems, Probability & Statistics, Machine Learning, Big Data, Advanced Python for Data Science
- 2. B.S. in Computer Information Systems, Data Science @ Edgewood University** AUG 2021 - MAY 2024
 - **GPA:** 3.6 (Completed in 3 yrs)
 - **Senior Award:** STEM Outstanding C.I.S. Student 2023-24
 - **Courses:** Data Analytics, Agile Methodologies, Discrete Mathematics, Financial & Managerial Accounting, C.S. Research

PROJECTS

- 1. Sports Replay Extraction Model @ New York University** FEB 2025 - MAY 2025
 - Developed a non-learning based **computer vision** model (OpenCV, OCR, CUDA) to extract sports replays from live broadcasts, achieving **92%** accuracy and a **2.6x** speed optimization.
- 2. Movie User Segmentation & Ratings Recommendation Model @ New York University** APRIL 2025 - MAY 2025
 - Tuned an implicit ALS recommendation model on a Spark/Hadoop cluster with 30M ratings and 330k users, using MinHash LSH for user clustering to improve accuracy by **170%** over the popularity baseline.

LEADERSHIP

- 1. Capstone Project Manager @ Edgewood University** JAN 2024 - MAY 2024
 - Led a team using Scrum to develop a Python **ETL pipeline** for PupilLabs eye-tracker data used in university psychology research.
- 2. Varsity Esports Team Captain @ Edgewood University** AUG 2021 - MAY 2024
 - Selected as **team captain** for **3** consecutive years, led strategy development, coordinated practice schedules, and mentored players, culminating in a Top 16 national ranking.

SKILLS

- 1. Programming Languages:** Python (**Libraries:** Pandas, NumPy, Matplotlib, Seaborn, Scikit-Learn, PyTorch, Cython, CUDA, OCR, OpenCV, BeautifulSoup, PyGame, Flask), SQL, PySpark, Java, PHP, R, JavaScript, HTML, CSS
- 2. Data & Statistical Analysis:** Power BI, Tableau, Excel, Office 365 Suite, Identifying trends and patterns, Regression analysis, Hypothesis testing, Predictive modeling, Data storytelling, Data Cleaning, Data Transformations, Data Visualizations, ETL Processes
- 3. Big Data & Machine Learning:** Hadoop, Spark, Dask, MapReduce, Google Cloud, AWS RedShift, Google BigQuery, Databricks, MongoDB, MySQL, NoSQL, Microsoft SQL Server Querying, Relational Databases, Partitioning, Database modelling (ER, EER Diagrams), database management, Computer Vision, Data Optimization
- 4. Other:** Git, GitHub, Jira, Notepad++, Scrum, Project Management, CI/CD, Computer Networking, IT Support, Academic Research, Finance, AI, Problem Solving, Written and Verbal Communication, Attention to Detail