

Lavya Midha

✉ lavyamidha015@gmail.com | 📞 413-800-1278
🌐 github.com/lavyam | 🌐 <https://lavyam.github.io/lavya-portfolio/>
🌐 [linkedin.com/in/lavya-midha-9ba408245](https://www.linkedin.com/in/lavya-midha-9ba408245)

Skills

Languages & Frameworks: Python, R, SQL, NoSQL, Rust, Shell Scripting, React, Git, Linux.

ML & AI: Supervised/Unsupervised Learning, NLP, LLMs (BERT, Hugging Face), Feature Engineering, Model Evaluation, TensorFlow, PyTorch, Scikit-learn, Keras, MLOps.

Data Analytics & Visualization: Excel (Advanced), Tableau, Power BI, Python Panel, R Shiny, Pandas, NumPy, Matplotlib.

Cloud & Big Data: AWS, Azure, GCP, Spark, Hadoop, Docker, API Integration, ETL Pipelines, HPC.

Work Experience

ReferU.AI, Lewes, Delaware

May 2025 – Present

Machine Learning Engineer

- Engineered a cross-state ad intelligence system by operationalizing demographic and court-level data, enabling precision-targeted legal campaigns across diverse jurisdictions.
- Designed a centralized Looker Studio dashboard integrating real-time Google Ads and Analytics data, enhancing executive visibility and accelerating campaign decision cycles.
- Developed an ML-based keyword optimization model that dynamically surfaced high-performing terms, increasing return on ad spend through automated bidding strategies.
- Strategic problem solving, machine learning, marketing analytics, API integration, executive dashboarding, cross-functional collaboration, Python, Applied ML, SQL, Looker Studio, Google Ads, Google Analytics, SEM Rush, GCP.

BU Spark!, Boston, Massachusetts

January 2024 - May 2024

Project Manager

- Led cross-functional teams on civic analytics initiatives, delivering data tools that informed public budgeting, voter engagement, and social impact strategy for nonprofit and government partners.
- Championed Agile implementation across engagements, improving delivery velocity by 40% and strengthening client alignment through structured feedback and iteration cycles.
- Tableau, Google BigQuery, Python, Agile project delivery, stakeholder engagement, cross-sector strategy.

BU School of Public Health, Boston, Massachusetts

February 2023 - December 2023

Data Science Research Assistant

- Led air quality analysis across 125+ public schools, delivering exposure models and time-series insights that shaped district infrastructure investment and health compliance strategy.
- Designed a computer vision pipeline to identify environmental disparities (e.g., window access), enabling equity-based prioritization of facility upgrades.
- Built a real-time monitoring dashboard for school operations, reducing reporting latency by 35% and enhancing responsiveness to temperature deviations.
- Python, SQL, R, Python Panel, scikit-learn, education infrastructure analytics, equity strategy.

Education

New York University

Expected May 2026

M.S. in Data Science (Industry Concentration)

GPA: 3.72/4

Relevant Coursework: Probability, Statistics, Programming, Intro to Data Science, Text as Data, Big Data, Machine Learning, Practical Training for Data Science.

Boston University

September 2021 - May 2024

B.S. in Data Science, Minor: Statistics

GPA: 3.75/4

Relevant Coursework: Data Visualization, Machine Learning and AI, Algorithms, Statistical Modeling, Probability, Big Data, Data Mechanics, Programming, Optimization, Regression Analysis, NLP, Biological Data Science, Linear Algebra, DEI in Tech, Civic Tech, DS for Good, Inference, Foundations of Data Science, Analysis of Variance, Calculus.

Leadership and Awards: Undergraduate Research Grant, Dean's List, Secretary – BU Data Science Association, International Peer Mentor, Peer Tutor, Admissions Ambassador.

Project Work

- Boston Police Overtime Strategy (RiseUp):** Led a forensic analysis of \$60M+ in overtime expenditures across units and ranks within the Boston Police Department. Identified structural inefficiencies and delivered interactive dashboards to advocacy partners, directly supporting budget reallocation and data-driven policy reform.
- COVID-19 Policy Benchmarking:** Built a cross-country analytics pipeline to standardize and compare pandemic response policies across 10 nations. Produced strategic dashboards and briefing materials that supported public health consulting engagements on containment effectiveness and preparedness planning.
- NYC Subway Sentiment Analysis:** Developed a BERT-based NLP pipeline to evaluate commuter sentiment from 10K+ geotagged tweets. Integrated findings with borough-level ridership data to uncover service satisfaction trends, informing mobility strategy for urban transit stakeholders.
- Personalized Recommendation Engine:** Designed a large-scale recommender and user segmentation system using PySpark for 330K users and 86K movies. Combined MinHash and ALS to surface behaviorally similar “movie twins,” outperforming baseline models on personalization accuracy and ranking relevance.