ANDREW SULLIVAN

sullivan.j.andrew@gmail.com
(781) 296 - 6009
Weston, MA
(Willing to relocate)
linkedin.com/in/andrewjosephsullivan
github.com/andrewsully

EDUCATION

Harvard University | John A. Paulson School of Engineering and Applied Sciences May 2024
MS in Data Science
- Relevant Coursework: Machine Learning, Deep Learning, Systems Development for CS, MLOps, and Data Visualization

Syracuse University | School of Information Studies May 2022
BS in Information Management & Technology (Concentrations: Data Analytics & Project Management)
- Summa Cum Laude, GPA: 3.94/4.00
- Awards & Activities: SU Scholars Nominee; *Mentor*, InclusiveU; *Runner*, Boston Marathon; *Philanthropy Chair*, Psi Upsilon

WORK EXPERIENCE

Data Scientist – End-to-End Model Development & Signal Processing at Scale Oct. 2024 – Aug. 2025
Titan Advanced Energy Solutions | Salem, MA
- Solely led the design, development, and delivery of a real-time failure prediction system for lithium-ion batteries as part of a DOE-funded SBIR Phase II project, authoring the 30+ page technical report submitted to the Department of Energy.
- Built an anomaly detection pipeline in Python using high-frequency ultrasonic data stored in AWS S3, applying moving averages and rolling-window statistics to capture evolving signal variance.
- Engineered 14 statistical, spectral, and transform-domain features to enable robust, generalizable failure detection—eliminating the need for manual tuning across diverse conditions.
- Designed a multi-layer validation framework combining signal smoothing, adaptive baseline comparisons, and cross-sensor consensus to flag anomalies while minimizing false positives.

Data Science Intern – Scalable Visualization & ETL Pipeline Development May 2023 – Aug. 2023
Invaio Sciences | Cambridge, MA
- Led the development of an advanced, scalable D3.js heatmap visualization tool, surpassing existing Looker capabilities, which facilitated more detailed molecular plate analysis and supported scientists in trial evaluations.
- Identified and resolved two critical issues affecting approximately 50% of custom visualizations, ensuring accuracy and functionality in data representation across company dashboards.
- Engineered complex Python ETL scripts to apply internal business logic to diverse data sources, facilitating streamlined loading into a centralized data lake.

Data Analytics Intern – Predictive Modeling & Strategic Business Insights Apr. 2021 – July 2021
Hi Marley | Boston, MA
- Formulated and presented four strategic business recommendations to C-suite executives projected to elevate customer satisfaction by over 20%, derived from predictive model analysis of more than 2,000 claim cases using R.
- Spearheaded the development of a new SQL-based application feature that allowed service leadership to track employee performance via an analytical dashboard, enhancing operational efficiency.

PROJECTS

Race for the White House (Media Bias Tracker) [*see here*] - *Data Viz. & MLOps, Harvard University* Fall 2022
- Collaborated on an award-winning full-stack web project recognized in the CS 171 Data Visualization Hall of Fame, using a fine-tuned RoBERTa (SiEBERT) model to analyze sentiment across 50,000+ TV news mentions of 2024 U.S. presidential candidates.
- Engineered a custom data pipeline to scrape and process closed caption text from national broadcasts, powering dynamic NLP-driven sentiment insights that revealed partisan media patterns.
- Designed and deployed interactive D3.js visualizations on Google Kubernetes Engine (GKE), enabling users to explore sentiment trends and media volume by network, time, and candidate.

URL-based Cyber Threat Prediction using Machine Learning [*see here*] - *Data Science 1, Harvard University* Fall 2022
- Achieved 95% accuracy classifying malicious URLs by engineering 21 predictive features, benchmarking five models, and selecting Random Forest based on cross-validated performance.
- Preprocessed and transformed raw URL data using Python, Pandas, NumPy, and scikit-learn, applying label encoding and feature scaling to optimize model performance.

SKILLS

Programming Languages: Python, R, SQL, JavaScript | **Data Science & ML:** Pandas, NumPy, Scikit-Learn, Tensorflow, Machine Learning, Probability | **Visualization:** Matplotlib, Seaborn, Plotly, D3.js, Tableau | **Cloud & DevOps:** AWS (S3), Git & GitHub, CI/CD, Testing / TDD