

James Lederman

Data science leader with a passion for technology and innovative data solutions

CONTACT

www.linkedin.com/in/james-lederman/

SKILLS

Tech Stack:

- Python
- R/RStudio
- SQL/RDBMS
- Scikit-Learn
- TensorFlow, PyTorch
- SpaCy, Gensim, NLTK
- Databricks, Spark
- Docker, Airflow
- AWS, GCP
- Git/Github
- Linux/Bash

Specializations:

- Machine Learning Algorithms
- MLOps, DevOps
- Runtime optimization
- Containerization and Virtualization
- NLP: NER, LLM, topic modeling
- Predictive Modeling
- Geospatial (GIS)
- Software development
- Agile mindset

EDUCATION

PhD, Neuroscience 2014
Stony Brook University
Stony Brook, NY

BA, Biology 2005
University of Pennsylvania
Philadelphia, PA

EXPERIENCE

Senior Data Scientist | Bigtincan - Waltham, MA/Remote

Managed a cross-functional team including data scientists, software engineers, speech and language experts, and product managers to develop an ML-based multilingual speech coaching and conversational AI software package.

Key accomplishments:

- Implemented automatic unbiased feature extraction (using wavelet transform and MFCC) on audio waveforms to improve model performance by ~20%
- Built multiple linear regression models to score sentiment of speech across multiple emotional dimensions (eg: confidence, engagement, persuasiveness)
- Built text classification model on sales-call transcripts to categorize call types for more effective sales-process monitoring
- Built predictive models to identify “will-do” statements in sales calls and recommend personalized action-items
- Worked closely with product managers to translate business problems into ML modeling problems and communicate complex ideas to key stakeholders to ensure business adoption

Data Scientist | AdTheorent New York, NY

Developed ML models, built scalable ML pipelines from data integration to post-deployment monitoring, and conducted predictive analytics for digital advertising campaigns. Key accomplishments:

- Built ML models for smart-bidding optimization, conversion/incremental conversion modeling, visitation and geospatial modeling, cross-device attribution using identity graph
- Developed a RESTful API for automation of lookalike models
- Collaborated with a team of 6 using an Agile framework to develop of fully documented Python modules and packages in accordance with CI/CD, DevOps, and PEP8 best practices
- Built and modified multiple end-to-end large scale (dataset size > 100 TB) automated machine learning pipelines using Airflow, Databricks, and Spark from data ingestion to post-production monitoring and reporting of model performance and data drift
- Served as information security “ambassador” - worked closely with Infosec team to implement best practices in data security and compliance

Data Science Fellow | The Data Incubator - New York, NY

- Predicted crop yields using satellite image data and deep learning models
- Created Tableau dashboard with interactive plots for visualization of results

Postdoctoral Researcher | Albert Einstein College of Medicine - New York, NY

- Designed and conducted experiments to test hypotheses about the contribution of specific neural circuitry to addiction behavior
- Wrote a custom R package to clean and integrate and visualize multimodal neural signal, video, and event data and apply statistical modeling techniques including cluster analysis, PCA/FA, and multiple linear regression to find correlation and infer causation of neural activity to animal motor behaviors.