
MACHINE LEARNING | ARTIFICIAL INTELLIGENCE | DATA ANALYSIS | STATISTICAL MODELING

Insight-driven data scientist with a strong foundation in computer science and an MS in Data Science from Tufts University. Skilled in machine learning, algorithm design, and large-scale data analysis. Built predictive tools that streamlined workflows and saved 100+ labor hours annually across a 17-person team.

EDUCATION

Master of Science in Data Science, *Tufts University, Medford, MA*

- GPA: 4.0/4.0
- Capstone Project: Enhancing ML-Based Galactic Photometric Redshift Estimation via Morphological Feature Integration
- Relevant Coursework: Python for Data Science, Reinforcement Learning, Artificial Intelligence
- Course Assistant for Introduction to Computer Science, Data Structures, and Introduction to Machine Learning.

Certificate in Data Science Fundamentals with Python and SQL, *IBM Skills Network*

Bachelor of Arts in Chemistry, *Bowdoin College, Brunswick, ME*

PROJECTS

Dynamic HTS Analytics for Drug Discovery

- Designed ML-driven analytics architecture using Python and Spark to optimize HTS workflows and reduce bottlenecks in drug discovery pipelines; evaluated approach on simulated datasets.

TF-IDF Sentiment Classification

- Classified sentiment in IMDb, Amazon, and Yelp reviews using TF-IDF features with scikit-learn; achieved 99.6% accuracy and 89% AUROC via GridSearchCV on MLP, SVM, and logistic models.

Monte Carlo Control for Racetrack Navigation

- Trained RL agent with off-policy Monte Carlo methods using NumPy to optimize pathfinding in penalty-based grid environments over 500k episodes.

Logistic Regression for Image Classification

- Achieved 3.75% test error using TensorFlow and scikit-learn by engineering models on texture and intensity features, with augmentation and hyperparameter tuning.

Transfer Learning for Grid Navigation

- Trained DQN agent using PyTorch with curriculum learning to achieve 97.7% success on novel scenarios ($p < 0.0001$) in obstacle-rich grids.

On- vs. Off-Policy RL Comparison

- Implemented and benchmarked off-policy and on-policy RL algorithms using custom simulation environments with NumPy.

RL Strategy Comparison in Bandits

- Compared epsilon-greedy vs. constant step-size strategies in non-stationary bandits; validated across 2000 runs using SciPy and visualized outcomes with Seaborn.

RELEVANT EXPERIENCE

Institutional Data Scientist / Instructional Leader for Mathematics, *Milken Community Schools, Los Angeles, CA*

- Automated performance tracking systems using Python and Pandas to monitor instructional impact across 17 educators.
- Built logistic regression and decision tree models to identify students at risk of not meeting grade-level proficiency, contributing to a 62% reduction in remediation cases over two years.
- Analyzed academic and operational datasets using SQL and Jupyter to optimize staffing and scheduling, resulting in a 15% improvement in instructional coverage efficiency across departments.

Senior Consultant, *Booz Allen Hamilton, San Diego, CA*

- Improved logistics response time by 12% and cut costs by delivering statistical models for defense resource planning.
- Recognized as a top-performing consultant for translating technical insights to senior government stakeholders.

SKILLS

Programming Languages Python, C++, SQL, Java, MATLAB

Tools & Platforms Docker, Excel, Git, Google Colab, Jupyter Notebook, LaTeX, VS Code

Frameworks & Libraries LangChain, LlamaIndex, Matplotlib, NumPy, pandas, PyTorch, scikit-learn, SciPy, Scrapy, Seaborn, Statsmodels, TensorFlow, Transformers, XGBoost

Big Data Tools Hadoop, Spark

ADDITIONAL EXPERIENCE

Department Chair for Mathematics, *Ojai Valley School, Ojai, CA*

- Improved math benchmark scores by 3% over two years with targeted, data-informed instruction.

AP Statistics Teacher, *Cushing Academy, Ashburnham, MA*

- Raised AP Statistics pass rates by 20% by designing and teaching a project-based statistics curriculum.

AP Chemistry Teacher, *Milton Academy, Milton, MA*

- Increased AP Chemistry exam scores by 15% over four years by designing/implementing data-rich lab experiences.