Mark Leschinsky

201-615-4641 | mark.leschinsky@stanford.edu

Education

Stanford University Stanford, CA

Bachelor of Science in Data Science | GPA: 3.73/4.00 *June 2027 (Expected)*

- **Relevant Coursework:** Applied Matrix Theory, Probability Theory, Statistical Inference, Programming Abstractions, *(Stochastic Processes (Master's Level), Financial Statistics (PhD Level) before summer 2026.)*

Certifications

Bloomberg Market Concepts

Bloomberg | View credential

Experience

Machine Learning Engineer Intern June 2025 – August 2025

CorInnova Cardiac Assist Technologies *San Francisco, CA*

- Developed an end-to-end supervised ML pipeline in Python (pandas, scikit-learn) to classify C-/C+/D patient heart failure from >150 morphological features (EDV/ESV) of 95 patients, achieving 87% accuracy and 0.72 F1-macro.
- Engineered preprocessing workflows with imputation, scaling, one-hot encoding, and hyperparameter tuning using GridSearchCV across KNN and Logistic Regression, reducing manual doctor review times at the company by 60%.
- Applied PCA and univariate tests to compress >150 variables into 3 key predictors while preserving 82% accuracy, reducing computational cost by 80% and enabling faster deployment.

Course Assistant, Principles of Data Science June 2025 – August 2025

Stanford University *Stanford, CA*

- Led 13 weekly discussion sections for 16 students covering pandas DataFrame ops, KNN/logistic/SVM modeling, K-Means clustering, hyperparameter tuning, and RMSE/precision-recall evaluation in scikit-learn.
- Graded 110 Colab notebook labs covering TF-IDF vectorization, dummy encoding, normalization, Euclidean/Manhattan distance metrics, cross-validation, and multivariate scatterplots.
- Hosted 7 office-hour sessions guiding students through real-world project scoping, Colab setup, unit testing & error handling, pipeline validation, and best practices.

Projects

Federal Lodging Rates | *pandas, scikit-learn, BeautifulSoup, Plotly, matplotlib* January 2025 – March 2025

- Compared federal per-diem rates with an engineered cost-of-living model and found average 25.7% underpayment for 3 million federal employees on work trips; presented research at the Stanford Data Science Capstone.
- Engineered a rate-limited scraper to query per-diem JSON endpoints for 26,321 ZIP codes, flattened nested dictionaries into pandas DataFrames, and merged with Census and Zillow features.
- Constructed a scikit-learn Pipeline with a ColumnTransformer (median imputation + StandardScaler for numeric; OneHotEncoder for categorical) feeding into KNeighborsClassifier; tuned params via GridSearchCV (macro-F1).

Technical Skills

Languages & Tools: Python, C++, Google Colab, Git
Data Processing & APIs: pandas, NumPy, requests, BeautifulSoup, JSON
Machine Learning: scikit-learn (GridSearchCV, ColumnTransformer, KNeighbors, RandomForest, GradientBoost)
Visualization: matplotlib, Plotly, seaborn

Extras

- Founded and formerly directed the world's largest student film festival—raised $322,000 from Google, led 1,594 volunteers, reached 795,381 students in 161 countries, and distributed a film on Prime Video to 100 million people.
- Drone photography enthusiast—my aerial shots of Stanford ended up on Wikipedia.
- Detail-oriented to a fault: I've corrected grammar in both *The Stanford Daily* and *The New York Times*.