Anthony J. Young, PhD

Data Scientist

(954) 732-6233 | anthonyjamesyoung@outlook.com

anthonyjyoung.com | LinkedIn: anthony-j-young | GitHub: youngant

SUMMARY

A data scientist with a passion for solving quantitative problems and working with large datasets, I am looking to leverage my 8 years of experience working with statistics and machine learning tools in astronomy to tackle real-world problems. I am eager to contribute my skills and expertise to a team dedicated to developing innovative solutions that drive value and generate actionable results for stakeholders.

SKILLS

Computer Languages	Python, SQL, Javascript (D3.js), HTML, CSS,
Python Libraries	Numpy, Pandas, Scikit-learn, Scipy, Matplotlib, Plotly, Tensorflow, Keras, NLTK
Machine Learning	Linear regression, Regularized regression, Multi-class classification, Natural language processing, Numerical optimization, Cox regression
Quantitative	Statistics, Hierarchical Bayesian inference, Markov Chain Monte Carlo
OS & Platforms	Bash, Git/GitHub, Docker, Azure, DataBricks, Linux/MacOS/Windows, Jupyter
Other	Data visualization, Technical writing, Public speaking, Object-oriented design

PROJECTS

Brain Cancer Survival Analysis — Erdos Data Science Boot Camp   April 2024—May 2024

- Analyzed large dataset of brain cancer case listings from a large government-maintained database (SEER).
- Built a custom data preprocessing pipeline to automatically handle missing data and problematic features.
- Trained a model using regularized Cox regression to predict risks associated with patient features, finding predictions of risk factors to be consistent with established clinical outcomes.

Reconstructing Gravitationally Lensed Galaxy Images with Machine Learning   Sept 2018—Jan 2024

- Developed a new linear regression-based Bayesian modeling scheme for recovering un-distorted "3D images" of gravitationally-lensed galaxies, improving results with low-quality data.
- Applied new method on real and simulated data, achieving higher accuracy and resulting in a 10% increase in peak signal-to-noise ratio with more uniform properties compared to previous state-of-the-art methods.

Building Large Galaxy Survey Results Database   Aug 2019—May 2020

- Processed large volumes of raw telescope data to measure galaxy properties as part of an international team.
- Created ETL process in an unfamiliar language for a previously-tedious manual data processing workflow.
- Analyzed observations and contributed to telescope observing proposals and peer-reviewed publications.

EXPERIENCE

Rutgers University   May 2017—Jan 2024
Graduate Research Assistant   *New Brunswick, NJ*

- Worked independently to develop solutions to assigned problems as a part of thesis research.
- Developed an extensible, maintainable Python software package to run a custom statistical analysis pipeline.
- Collaborated within an international team to analyze large dataset and report results for multiple projects.
- Published numerous peer-reviewed research articles and presented results at large national conferences.
- Mentored 5 undergraduate students with minimal prior experience, resulting in successful presentations of their research to program leaders.

EDUCATION

Rutgers University, PhD, Physics and Astronomy   Jan 2024
Florida State University, BS, Physics & Applied Mathematics   May 2016