

Zach Yek

Data Analyst/Scientist

Contact

716-680-4665

✉ yek9367@fredonia.edu

📍 Fredonia, NY, USA

🌐 [linkedin.com/in/zachtheyek](https://www.linkedin.com/in/zachtheyek)

🐙 github.com/zachtheyek

🌐 zachtheyek.github.io

Education

B.S. Theoretical Physics

SUNY Fredonia

Aug 2019 - May 2023

GPA: 3.7

Minors: Computer Science,
Mathematics

Relevant coursework:

- multivariable calculus
- linear algebra
- probability & statistics
- machine learning
- digital image processing & computer vision
- data structures

Skills

- Python (NumPy, Pandas, PyTorch, TensorFlow, Keras, XGBoost, Matplotlib, Seaborn)
- SQL (MySQL, BigQuery)
- Git
- Linux
- Cloud Computing (GCP)
- Deep Learning
- NLP

Featured Experiences

2021 - 2022

Berkeley SETI Research Center, UC Berkeley | *Berkeley, CA*

Research Assistant

Q: The search for extraterrestrial intelligence is often limited by local interference

- Scraped and cleaned 95,000 radio observations using **BigQuery** and **Pandas**
- Developed a modified **Transformer** BERT model in **PyTorch**
- **Parallelized training** across 16 GPUs on a **GCP** VM instance
- Achieved 89.66% **classification** accuracy using **k-fold cross-validation**

<https://github.com/zachtheyek/autoSETI>

2019 - present

Department of Physics, SUNY Fredonia | *Fredonia, NY*

Undergraduate Researcher

Q: Protostellar astronomy typically requires multi-wavelength observations

- Physically simulated protostellar systems using **radiative transfer models**
- Derived a linear correlation (**correlation coefficient** 0.95) at 73.6 microns
- Demonstrated that a single observation at micrometer wavelengths is statistically sufficient, and provided a framework for computing uncertainties
- Cut down observing time by 190 hours, and observing costs by \$390,000 (per target object)

<https://github.com/zachtheyek/Protostellar-Luminosity>

Q: Multiple competing hypotheses exist for a rare stellar phenomenon

- Synthesized 100 distinct observations of a target star into a continuum image
- Designed custom **Python** scripts to enhance image quality by a factor of 248
- Implemented **API shell scripts** to extract a stellar disk mass of 0.024 solar masses, implying a disk-to-star mass ratio of 5%
- Interpreted our results to be in direct contradiction with several existing hypotheses, narrowing down the possible explanation space by ~20%

<https://github.com/zachtheyek/HBC722>

Q: CCD cameras become inaccurate when oversaturated at long exposure times

- Outlined a procedure to acquire 30 images at varying exposure times
- Established a baseline for image quality using the data at low exposures
- Applied **linear regression** to extrapolate the baseline out to higher exposures
- Determined an optimal correction factor of 0.1% to 1.8% using **polynomial regression**, with a **reduced-chi squared** value of 0.309

<https://github.com/zachtheyek/CCD-Linearity>

Selected Projects

- **Recommendation Engine:** Scraped data using the myanimelist.net API. Developed a pairwise similarity recommendation framework with **NumPy** and **Pandas**. Optimized performance for chained queries.

<https://github.com/zachtheyek/Anime-Rec-Bot>