256.585.7439
ychen.adastra@gmail.com

Yu Chen

New York, NY

ychenxastra.github.io
LinkedIn: yu-chen-astra

EDUCATION

Doctor of Philosophy, Space Science, The University of Alabama in Huntsville **June 2018 — Dec 2020**
Master of Science, Space Science, The University of Alabama in Huntsville **Aug 2015 — May 2018**
Bachelor of Science, Atmospheric Science, Nanjing Univ. of Info. Sci. & Tech. **Sep 2011 — June 2015**

EXPERIENCE

Center for Space Plasma and Aeronomic Research **Huntsville, AL**
Research Scientist Jan 2023 — present
Postdoctoral Researcher Jan 2021 — Dec 2022

- Utilized Python to process and analyze over 50GB of high-resolution raw datasets over 50 years. Designed algorithms to identify and catalog typical events, establishing a local database with over 140,000 entries.
- Analyzed over 140k events, extracted 60+ characteristics per entry, performed time-series and statistical analyses, deriving insights to enhance understanding of phenomena and propose hypotheses.
- Developed an open-source Python package for event studies, automating characteristic extraction and visualization, reducing manual input and user workload from months to minutes.
- Employed MATLAB and Matplotlib for data product visualization, summarized findings in 20 peer-reviewed journals and presented at major conferences for 18 times, effectively communicating to diverse audiences.
- Led two national research grants as Principal Investigator and collaborated on multi-million-dollar projects, mentored students in an NSF-funded program, contributing to research and educational development.

PROJECTS

PyGS: a Python package for solar wind structure analysis Python, Numpy, Pandas, Scipy, Matplotlib

- Optimized detection algorithm with advanced analytics and theoretical concepts, migrated separate Matlab-based techniques, refactored and merged 30+ scripts into a Python library, reducing manual data processing and script execution efforts.
- Upgraded data cleaning modules and streamlined system architecture, significantly reducing time and space complexity, leading to a 100x performance increase and broadened analytical applicability.
- Released a Python package on GitHub, meeting Heliophysics community standards, provided detailed documentation and tutorials, and promoted its use in the scientific community.

Database of small-scale magnetic flux rope Python, Matlab, Excel, PostgreSQL, HTML

- Employed Python to extract events from large spacecraft datasets over 50 GB. Developed sophisticated filtering algorithms to process over 200k candidates, forming the foundation for creating a local database.
- Managed the database using PostgreSQL, aggregating selected events in an online platform. Maintained and updated the online database via HTML to facilitate efficient and convenient user inquiries.
- Conducted in-depth analyses such as linear regression, to characterize product properties. Provided parameters in various formats, assisting users in understanding and extracting relevant information for their research interests.

Analyses of Football Matches and Predicting Results Python, Pandas, Scikit-learn, Seaborn

- Analyzed 20+ years of English Premier League data, focusing on Manchester United, and visualized statistics to identify key and controversial factors impacting team performance.
- Developed machine learning pipelines with models like Naive Bayes and Random Forest to predict match outcomes, selecting the best model based on performance metrics and applicability.
- Enhanced model accuracy using techniques like rolling averages, feature selection, and PCA. Integrated findings with real-world scenarios to explore causes of suboptimal performance.

TECHNICAL SKILLS

Programming Python (Numpy, Pandas, Scipy, Matplotlib, Scikit-Learn, Seaborn), MATLAB, C/C++
Tools & Software SQL (MySQL & PostgreSQL), Latex, Git, Jupyter Notebook, Excel
Platforms & OS Linux, MacOS, GitHub