256.585.7439
ychen.adastra@gmail.com

Yu Chen

New York, NY

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

SUMMARY

- Seasoned quantitative PhD in large-scale data analysis including processing, visualization, and algorithm design.
- Manipulated data and deduced mission-critical insights by applying statistical computer languages such as Python and SQL.
- Experienced in leadership, mentoring interns, and collaborating with cross-functional teams on multiple projects.

TECHNICAL SKILLS

Programming	Python, MATLAB, C/C++ | Predictive Modeling, Hypotheses testing, Machine Learning
Database Processing	SQL (MySQL & PostgreSQL), AWS (EC2, S3, & RDS), Google Cloud BigQuery, PySpark
Data Analytics	A/B Testing, Casual Inference, Tableau, Google Analytics, MS Excel

EXPERIENCE

Center for Space Plasma and Aeronomic Research — **Huntsville, AL**

Research Scientist — Jan 2023 – present

Postdoctoral Researcher — Jan 2021 – Dec 2022

- Developed ETL pipelines using Python to process and analyze 100+ million data points in multiple dimensions, identifying aurora-related physical phenomena from spacecraft sequential data patterns and improving efficiency by 50%+.
- Established databases for physical phenomena with 200k+ entries using PostgreSQL, PySpark, and HTML, equipped with 100+ parameters based on user needs, reducing processing time by 25% and facilitating inquiries.
- Identified 10+ key factors affecting physical phenomena using statistical methods like ANOVA and linear regression, improving both identification and prediction of their variations with phenomena such as the solar cycle.
- Created Tableau dashboards to track relationships between key factors like location and physical phenomena in real time, and provided weekly reports to the team, enabling easy data-driven insights and efficient communication.
- Built and trained machine learning models (random forest, XGBoost, and decision trees) on data points from 20+ daily sources, and fine-tuned models using multiple techniques such as PCA, improving the model accuracy by 15%.
- Led 2 teams, mentored interns (project design, tool usage, and provided weekly feedback), coordinated the CUWiP, and collaborated with cross-functional teams (30+ people) on multi-million-dollar projects.

PROJECTS

Open-source Python Package for Phenomena Analysis — Python, Numpy, Pandas, Scipy, Matplotlib

- Designed and optimized algorithms for in-depth analysis of phenomena from sequential data and delivered Level III data products equipped with intuitive visualization, empowering users to extract relevant insights.
- Upgraded existing ETL modules and streamlined system architecture as well as reduced time and space complexity, leading to a 100x performance increase and broadened analytical applicability.
- Refactored over 30 Python scripts of algorithms and released a package containing 6,500+ lines of code on GitHub.

Analysis of English Premier League Matches and Prediction — Python, Pandas, Scikit-learn, Seaborn

- Processed and cleansed data from over 20 years of the English Premier League (11,000+ rows) and performed EDA to analyze the impact of controversial factors on match outcomes, extracting valuable insights through data science techniques.
- Forecasted soccer match outcomes using a random forest model, achieving an AUC of 0.87 for predicting home team wins.
- Improved prediction performance by 10% via 5-fold cross-validation and GridSearchCV.

A/B Testing Outcomes of E-commerce Company — Python, Pandas, Numpy, Scipy, Seaborn

- Analyzed A/B testing results in hypothesis testing and extracted causality between UI change and conversion rate.
- Performed the sanity check to discover the imbalanced exposure and automated data quality check, ensuring the accuracy and reliability of the test implementation.
- Investigated relationships between conversion rate and multiple factors (geographic locations, time, etc.) to optimize the targeted marketing campaigns and conducted Z-Test to provide recommendation based on long-term consideration.

EDUCATION

Doctor of Philosophy, Space Science (STEM major), The University of Alabama in Huntsville — **Jun 2018 – Dec 2020**

Master of Science, Space Science (STEM major), The University of Alabama in Huntsville — **Aug 2015 – May 2018**

Bachelor of Science, Atmospheric Science, Nanjing Univ. of Info. Sci. & Tech. — **Sep 2011 – Jun 2015**