

SKILLS

MACHINE LEARNING: classification, regression, clustering, forecasting, feature engineering

SOFTWARE AND PROGRAMMING

LANGUAGE: Python (scikit-learn, numpy, scipy, pandas, pyspark, keras, dask, numba), SQL, Hadoop, Linux, Git, C++

SUMMARY

Dedicated and trained data scientist with a background in computational astrophysics for 10+ years: experienced in code development for numerical simulation in supercomputer clusters and the statistical pipeline for data analysis. I have developed state-of-the-art simulation models and python analysis tools to understand the nature of black holes. With my expertise in coding and statistical analysis, as a Junior Data Scientist at Veda, I am passionate to develop machine learning algorithms that help the company gain insight into catching hidden messages or trends from datasets in business.

EMPLOYMENT

UNIVERSITY OF AMSTERDAM

Postdoctoral Researcher

Amsterdam, Netherlands

Sep. 2018 to Mar. 2022

- Developed and performed a GPU-enabled parallelized simulation code in supercomputer clusters, which resulted in >10 times faster than previous CPU codes
- Developed python visualization/analysis tools to produce synthetic black hole images from large scale simulation datasets (>1TB)
- Co-advised three Ph.D. students and trained one master's student at the University of Amsterdam to help them initiate dissertations

SHANGHAI ASTRONOMICAL OBSERVATORY

PIFI Research Fellow

Shanghai, China

Nov. 2015 to May 2018

- Developed a CPU-parallelized algorithm and python analysis tools to evaluate theoretical models of black hole feedback
- Awarded an AAS and IOP Publishing China one of Top 1% Cited Papers Award in the Astrophysical Journal

UNIVERSITY OF WISCONSIN-MADISON

Research Assistant

Wisconsin, USA

Sep. 2009 to Aug. 2015

- Designed and tested theoretical black hole jet models, which resulted in 4 research publications
- Developed python tools to analyze the statistics of the simulated jet features

PROJECTS

WEB TRAFFIC TIME SERIES FORECASTING

Jan. 2023

- Aimed to predict future web traffic for approximately a total of 145k Wikipedia articles to make better traffic control decisions
- Conducted feature engineering of time series data and built/evaluated three different forecasting models
- Found that the daily traffic had a strong weekly seasonality and predicted the next 60 days of traffic successfully by the LSTM model

WEST NILE VIRUS PREDICTION

Dec. 2022

- Aimed to build a model that predicts outbreaks of West Nile virus in mosquitos using the given environmental conditions, such as weather, season, and location in the City of Chicago
- Managed highly imbalanced data by undersampling technique and solved binary classification problem by developing five different models
- Predicted the presence of viruses successfully using the Random Forest model and found that wet and warm air could promote a favored environment for mosquito proliferation and thus the virus epidemic

EDUCATION

Springboard Data Science Career Track

Feb. 2023

Certification 2023

8-month intensive course in data science, machine learning, Python and SQL.

Portfolio Projects: <https://astrodoo.github.io/portfolio/>

University of Wisconsin-Madison

Aug. 2015

PhD, Astrophysics

Seoul National University

Dec. 2008

Master of Science, Astrophysics

Seoul National University

Dec. 2005

Bachelor of Science, Physics and Astronomy