Ying Li

https://yingli9718.github.io/ yingli.atmos@gmail.com
www.linkedin.com/in/yingli-atmos/ (609)-356-6330
US Permanent Resident Vienna, VA

Highlights

- 10+ years experience in Machine Learning, data analysis, statical analysis, visualization and climate modeling
- Strong writing and verbal communication skills with 20+ papers in top journals and 50+ conference presentations
- Good collaboration skills with many global collaborators on different projects
- Author of analyses and visualizations of an upcoming textbook in climate data analysis
- AWS Certified Cloud Practitioner, Machine Learning (Coursera), Data Scientist with Python (26 courses on Datacamp)

Experience

- **Colorado State University**
  Research Scientist II *Aug. 2016 - Present*
  - Led a 4-yr project funded by NSF and published 8 lead author papers in top academic journals
  - Successfully brought in ∼$1,000,000 in funding from NSF and NASA by leading projects as principal investigator
  - Analyzed and diagnosed periodic patterns in data from observations and idealized oscillator models using techniques including regression, principal component analysis (PCA), power spectral analysis, time and space filtering. Discovered a statistically robust periodic behavior in time series that has been used for climate prediction
  - Designed and performed test and control numerical experiments with climate models to investigate the impact of clouds on climate patterns. Provided the insights for improving model performance
  - Mentored one PhD student and served as thesis committee member
- *Postdoctoral Scientist* *Jan. 2012 - Jul. 2016*
  - Leveraged Machine Learning to quantify the magnitude of the contributing factors on the clouds formation and variation from the raw ∼400 millions satellite profile data. Provided a benchmark analysis for model performance evaluation and model parameter tuning
  - Performed the large ensemble of test and control numerical experiments with high-resolution global climate model on high-performance computing clusters at National Center for Atmospheric Research (NCAR). Improved understanding of the role of the ocean circulation in climate variability
  - Quantified and visualized the spread in circulation pattern responses to same global warming forcing in six different climate models. Provided the uncertainty attributions to climate change simulation
- **Princeton University**
  Research Assistant *Sep. 2006 - Nov. 2011*
  - Investigated the casual relationships between the circulation patterns using techniques including regression, time series analysis, PCA, based on the 2000-yr, high-dimensional, tera-byte scale data from numerical simulations with Earth system models. Provided comprehensive picture of the global teleconnectivity
  - Designed and performed test and control numerical experiments running on high-performance computing clusters at Geophysical Fluid Dynamics Laboratory (GFDL). Provide evidences on the relative roles of surface boundary conditions over the particular sectors in the globe

Side Projects see more at https://yingli9718.github.io/side_project.html

- Scraped data from Airbnb website and built Machine Learning models to predict the booking rate and price, provide recommendation list to improve the listing, and benchmark analysis for listing hosts to maximize their revenue
- Built a COVID-19 dashboard to visualize and track the spread of the virus using Python and Tableau

Education

- **Princeton University** *Ph.D. in Atmospheric and Oceanic Sciences*
- **Zhejiang University** *B.S. in Atmospheric Sciences*

Technical Skills

- **Programing tools**: Python (scikit-learn, NumPy, Scipy, Pandas, TensorFlow, Keras, Plotly, BeautifulSoup), SQL, R, Shell scripting, Git, Fortran, C, PostgreSQL, MATLAB, HTML, CSS, LaTex
- **Machine Learning tools**: Linear/Logistic Regression, Classification, Decision Trees, GBM, XGBoost, Random Forest, Neural Networks, NLP
- **Statistics**: Regression, Bayesian and Monte Carlo methods, Time Series, PCA, A/B testing
- **Data Visualization tools**: Tableau, Matplotlib, Seaborn, Plotly
- **Big Data & Hadoop**: HDFS, MapReduce, Hive, Spark