PAUL S. LIN

3704 North Charles Street #1304 Baltimore, Maryland | 410-794-6079 | paullin3679@gmail.com

EDUCATION

University of Pennsylvania, School of Engineering and Applied Sciences, Philadelphia, PA **Dec 2021**
Candidate for Master of Science in Engineering in Computer and Information Science **GPA: 3.90/4.00**
- Coursework: Big Data Analytics, Applied Machine Learning, Database & Information Systems, Artificial Intelligence

University of Pennsylvania, Stuart Weitzman School of Design, Philadelphia, PA **Dec 2021**
Candidate for Certificate in Geospatial Information System (GIS) & Spatial Analysis **GPA: 4.00/4.00**
- Coursework: Geospatial Data Science, Geospatial Software Design, Remote Sensing, Advanced Topics in GIS

University of Pennsylvania, College of Arts and Sciences, Philadelphia, PA **May 2021**
Graduated summa cum laude, Bachelor of Arts in Earth Science with Honors **GPA: 3.91/4.00**
- Coursework: Cloud & Scalable Computing, Data Structures & Algorithms, Linear Algebra, Multi-Variable Calculus
- 2021 Udall Scholarship: for pursuing leadership, public service, and commitment to environmental issues
- 2021 Henry Darwin Rogers Awards: awarded to one graduating earth science senior major for excellence in study

SKILLS

Programming: Python, Java, JavaScript, SQL, Cypher, Swift, Linux, AWS (EC2, DynamoDB, RDS), Oracle/MySQL, MongoDB, neo4j, web (CSS/HTML, Bootstrap, Node.js/express, jQuery, Django, ZingChart)
Data Science: scikit-learn, NumPy, pandas, Matplotlib/seaborn, PyTorch, pyspark/Apache Spark/MLlib, DataBricks
Geospatial: ESRI (spatial analyst, ArcGIS Pro, ArcPy), remote sensing/Google Earth Engine, geopandas, cartopy, xarray
Foreign Language: Chinese, Korean

EXPEREINCE

Pharmaceutical Product Development (PPD), Wilmington, NC (Remote) **Jul 2021 – Present**
Intern-Graduate
- Developed time-series models for forecasting one-month-out volume and revenue predictions of lab supply and tests
- Analyzed historical data to evaluate item-revenue distribution and to identify the top revenue-generating items

Universities Space Research Association (USRA), Huntsville, AL (Remote) **Jun 2021 – Aug 2021**
Intern, Student
- Generated machine learning models for predicting hourly PM2.5 values from sensor data that reduced forecasting RMSE by eight-fold and percent error by 3-fold compared to study-based formula forecasting methods
- Implemented ETL pipeline for periodically collecting data from open-source databases, inputting processed data into backend machine learning models, updating entries within internal databases, and visualizing outputs onto maps
- Deployed Node.js web app in AWS for displaying PM2.5 observation and forecast values at U.S. embassy sites

University of Pennsylvania, Dept. Computer and Information Science, Philadelphia, PA **Jan 2021 – May 2021**
CIS 545 – Big Data Analytics, Teaching Assistant
- Graded assignments, hosted office hours, and led recitations for graduate data science class of over 350 students
- Course Topics: data wrangling/processing/modelling, machine learning & neural networks, parallelizable computing

NASA Goddard Space Flight Center, Greenbelt, MD **Jun 2019 – Aug 2019**
Global Modeling and Assimilation Office Intern
- Wrote Python scripts that processed various GEOS datasets and identified potentially erroneous data values
- Developed leaflet web interfaces that allowed users to display variations of maps showcasing error distributions

NASA Goddard Space Flight Center, Greenbelt, MD **Jun 2018 – Aug 2018**
Science Data System Intern
- Developed poster and seminar presentation that promoted for the usage of cloud computing among NASA scientists
- Documented the process for utilizing AWS virtual machines for processing earth science data and scripts

PROJECTS

Latitudinal Shifts of Plant Functional Types within the Great Plains
- Analyzed the spatiotemporal trends and environmental correlations of plant functional types within the Great Plains
- Utilized Python, ArcGIS, Google Earth Engine to process remote sensing data and build machine learning models

Cattell's 16 Personality Factors Test
- Developed machine learning models (clustering, random forest, regression) for predicting demographic characteristics (countries, gender, age) from personality types with pandas and scikit-learn

FoodFinder 3000
- Designed a web-interface application for providing restaurant and itinerary recommendations based on yelp reviews
- Utilized SQL and Oracle for database management and queries and Node.js, express, react for web architecture

PennBook
- Utilized AWS, DynamoDB, Node.js, & express to build a "mini-Facebook" application with the supported features of user profiles, walls, posts and comments, chats and messages, friends, newsfeed updates and recommendations