
Professional Experience

Research Consultant.....July 2021 to Current
Progl.ai, Baltimore, MD

- Study the extrapolative behavior of various ML algorithms implemented using **Python sklearn** package
- Build an web app to conduct human behavioral experiments using **HTML, Javascript and Flask**, and backend database management using **SQLAlchemy** hosted on heroku
- Prepare the NeurIPS 2021 workshop on inductive bias of machine extrapolative behavior

Research Intern May 2021 to July 2021
Microsoft Research, Redmond, WA

- Developed a web application that automates the process of causal inference using **React and Typescript** to power front-end and **python Flask** to service back-end in the context of human trafficking and COVID-19
- Built and designed an **end-to-end** pipeline of **causal inference** using **Python DoWhy** package to deploy a research application product branded as ShowWhy
- Coordinated a team of 6 engineers of different technical background to build the initial ShowWhy application

Data Science Intern January 2021 to May 2021
Johns Hopkins University School of Medicine, Baltimore, MD

- Developed a data preparation pipeline for JHU COVID-19 initiative using **probabilistic linkage** methods to provide backend database management in **R**
- Built a data analytic pipeline to parse hospital datasets via **causal inference** methods to provide a clinical data science guideline for COVID-19 vaccination distribution

Graduate Summer Research Intern.....July 2020 to Sep 2020
Johns Hopkins University School of Medicine, Baltimore, MD

- Implemented the **deep learning** framework for radon transformation used in CT image reconstruction known as iRadonMAP in **Matlab**
- Produced augmented image dataset from ImageNet to train iRandomMAP algorithm for training the network

Data Science Consultant January 2021 to May 2021
MindX, Bethesda, MD

- Built an analytical pipeline to parse out biometric signals from the **multivariate time-series** hologram signal datasets in **python** using **sklearn** and **tensorflow**
- Cleaned and pre-processed real-world multimodal datasets using **pandas** and **hyppo**

Post baccalaureate IRTA research fellowNovember 2018 to May 2021
The National Institutes of Health, Bethesda, MD

- Built a prediction pipeline for HIV detection by HIV antibody titer in **Python** using **pandas** and **sklearn**
- Conducted parametric/nonparametric multivariate linear regression analysis of the national omics datasets such as metabolomics and proteomics using **Python** and **R**

Graduate Student June 2017 to October 2018
Virginia Commonwealth University, Richmond, VA

- Built automated fourier transformed signal detection program in **Matlab** for mouse behavioral experiments

Recent Publications



1. *Interclass GPCR heteromerization affects localization and trafficking* (2020), [Science Signaling](#)
2. *Fully automated head-twitch detection system for the study of 5-HT2A receptor pharmacology in vivo* (2019), [Scientific Reports](#)

Education

Masters of Science and Engineering, Biomedical Engineering (GPA: 4.0/4.0) **May 2021**
Johns Hopkins University - Whiting School of Engineering, Baltimore, MD

Skills

Python (Numpy, Pandas, Sklearn, Flask, Pytorch, DoWhy), SQL, Git, R, Matlab, Typescript, React, HTML

 Baltimore, MD
 jshin.m@gmail.com

Jong M. Shin

 [linkedin.com/in/jshinm](https://www.linkedin.com/in/jshinm)
 github.com/jshinm

Professional Projects

Automated end-to-end causal inference application (MSR)

- Devised application design from a data scientist perspective and laid out the basic architecture from scratch
- Built example framework of the app using Fluent-UI, typescript, and python-flask
- Wrote libraries and notebooks of end-to-end data science pipelines to serve as the development plan

Inductive bias experiment (JOVO Lab) [[GitHub](#)]

- Implemented ML models from sklearn and trained on nonlinear simulation data
- Generated mathematically derived posterior probability for exclusive OR and spiral dataset
- Implemented point-wise Hellinger distance and explored extrapolative behavior of ML models

Web application for human behavioral experiment (JOVO Lab) [[GitHub](#)]

- Developed the web application for human behavioral experiment to collect inference performance
- Designed front-end using HTML/CSS/JavaScript and powered back-end using python (Flask, SQLAlchemy)

Probabilistic linkage and causal inference on COVID-19 (JHU CDEM) [[GitHub](#)]

- Built an automated system for probabilistic linkage pipeline to clean and join multiple hospital datasets
- Built a data analytic pipeline to parse hospital datasets via causal inference methods to provide analytical guideline for COVID-19 vaccination distribution within underrepresented minorities

Multivariate time-series hologram signal parsing (MindX) [[GitHub](#)]

- Cleaned and preprocessed proprietary hologram time-series datasets
- Investigated statistical significance of the signals detected from the datasets by conducting multivariate two-sample tests using in-house statistical software written in python