

# HUGH NGUYEN

[linkedin.com/in/hieuhughnguyen](https://www.linkedin.com/in/hieuhughnguyen) • [hieu-hugh-nguyen.github.io](https://github.com/hieu-hugh-nguyen) • [hnguye78@jh.edu](mailto:hnguye78@jh.edu) • +1 (860) 308-3839 • New York City

Scientist with 5+ years of experience in extracting insights and developing end-to-end algorithms from big messy biomedical data of various sources; 8+ years in medical research (cardiology, critical care, sleep, oncology); 2+ years of industry experience in big-tech and startups environments; expertise in AI/ML, statistics, causal inference, time-series, signal processing, and epidemiology. Strong research interest in ML method and analytical tool development for applications in medicine, biology, and health technologies.

## EXPERIENCE

---

### APPLE

Cupertino, CA

#### *Data Scientist PhD Co-op – Health Technologies*

Jun 2021 – Dec 2021

- Led 2 projects investigating users' patterns using 2+ million days of real-world tracking data from the Apple Watch and Apple Health App, leveraging methods in stats, AI/ML, causal inference, and time-series; resulted in 1 ML conference abstract, 2 first-authored internal papers, and helped guide strategic directions in the health space.
- Optimized 3 health product features for Apple Watch Series 8 and future gen, using insights extracted from data and end-users as the focus, in collaboration with 5 teams from software, design, and research.
- Led a team of 4 securing a spot in the finale of Apple's Shark-Tank-like internal competition (2% acceptance rate) to present ideas in front of Apple executives to improve user's eye protection (applicable to iPhones, iPads, and Macbooks).

### PERTHERA AI

Boston, MA

#### *Biomedical Data Scientist Intern – Computational Biology*

May 2020 – Dec 2020

- Built and deployed the company's first AI-powered interpretable outcome prediction algorithm for treatment response in patients with pancreatic cancer from Real World Evidence (RWE) multimodal data (genomics, proteomics, tumor profiling, imaging); listed as a co-inventor for a resulting patent. This work is actively being implemented by the company for other outcomes and datasets.
- Developed a smart literature recommendation system that recommended the most relevant papers to oncologists from 48,000+ Pubmed research papers using NLP methods.

### JOHNS HOPKINS HOSPITAL

Baltimore, MD

#### *Health Informatics Researcher – Precision Care Medicine*

Sep 2018 - Present

- Develop personalized real-time early warning models for 9 critical illnesses and injuries in the ICU, leveraging ML methods on various types of high-dimensional medical data (e.g., biosignals, EHR/EMR data, images, clinical trials, observational studies), led to the discovery of new biomarkers and insights for intervention and treatment.
- Led 8 research teams (40 ppl in total) on every step of their data science cycles to design, build, and ship models.
- Won Investigation Awards at RESS'19 & SCCM'20 – some of the most important cardiac arrest and critical care meetings worldwide.

### MEDTRONIC

North Haven, CT

#### *Engineer – Minimally Invasive Therapies*

Aug 2016 - May 2017

- Designed, analyzed, and prototyped a portable, easy-to-use testing fixture device for a surgical stapler used in laparoscopic surgery, reduced manufacturing costs and enhanced affordability by 500%; collaborated with different teams (product, business, engineer) to integrate ideas and communicated business recommendations to stakeholders.

## EDUCATION

---

### JOHNS HOPKINS UNIVERSITY, Ph.D. and M.S. in Biomedical Engineering – Data Science Track

Jan 2023 (Expected)

- Topic: ML Methods for Time-to-Failure Analysis Using Longitudinal Multimodal Data in Cardiovascular Diseases

### TRINITY COLLEGE, B.S. in Mechanical Engineering, President's Fellow, Full-Ride Scholarship

May 2017

## TECHNICAL SKILLS

---

- **Projects:** too many to list here, please visit <https://hieu-hugh-nguyen.github.io> for details
- **Machine Learning/Statistics:** Survival Analysis, Statistical ML (Forests, Boosting, SVMs), Deep Learning (RNNs, Autoencoders, CNNs, GANs, Transformers), Explainable AI, Time Series, Causal Inference, Signal Processing, NLP, Data Mining, Feature Selection/Engineering, PHC/Distributed/Cloud Computing, Model Optimization/Stacking/Deployment, Transfer Learning
- **Software Tools:** R, Python (TensorFlow, Keras, Pytorch, Sklearn) (6+ yrs experience), SQL (PostgreSQL, MySQL, BigQuery) (4+ yrs), Apache Spark, Linux/Unix Shell Scripting (4+ yrs), MATLAB (8+ yrs), Java, Bash, Git, Vitrea Imaging, SAS
- **Cloud platforms/Container technologies:** Google Colab, AWS, Microsoft Azure, Databricks; Docker, Kubernetes

## OTHER ACTIVITIES

---

Guest speaker/lecturer for 5 data science sessions that: i) taught data tools to 80+ physicians and healthcare professionals, ii) sparked interest in 'AI in Medicine' in 50+ students, and iii) offered advice to 200+ data science internship-seekers, sponsored by VinAI Research, YSEALI, FPT Software, and Johns Hopkins.