# James Wu

jswu18@gmail.com
linkedin.com/in/jswu18
github.com/jswu18
jswu18.github.io

## WORK EXPERIENCES

**UNIVERSITY COLLEGE LONDON** — London, UK
Research Assistant — *Oct 2022 - Present*
**Non-Parametric Bayesian Inference in Function Spaces** (GPyTorch, PyTorch)

- Developing a novel sample-based method to learn Bayesian posteriors directly on function spaces, avoiding variational approximations (see Github)
- Work in progress with Veit D. Wild & Jeremias Knoblauch for **JMLR 2024**

**Generalised Variational Inference for Gaussian Processes** (JAX, Flax, Optax)

- Designed and implemented a complete **Gaussian process** package from scratch for JAX compatibility (see GitHub)
- Proposed a new class of **variational inference** objectives to linearly scale for big datasets, addressing the challenges of existing methods that scale cubically

**Kernel-based Hypothesis Testing** (JAX)

- Implemented **integral probability metrics** such as the **maximum mean discrepancy** (MMD) and the **kernel Stein discrepancy** (KSD) (see GitHub)

**REVOLUT LTD.** — London, UK
Machine Learning Engineer (Junior → Mid → Senior) — *Aug 2019 - Sept 2022*
**Global Card Issuance for Company Profitability** (Pyro, PyTorch, BayesOpt)

- Strategically managed global card issuance through **portfolio optimisation** to earn interchange rebates exceeding $100M annually
- Constructed training and inference pipelines for **Gaussian processes** and **Bayesian optimisation** to forecast growth, spending, and activity for hundreds of granular cohorts

**Support Automation** (PyTorch, Scikit-Learn, Airflow)

- Designed models for semantic search and intent recognition for customer chat automation with **one-shot learning** and **transformer** sentence embeddings
- Improved embeddings with **multi-tasked learning** to effectively leverage limited corpora such as FAQs, pre-defined chat responses, and agent chat logs
- Applied **model distillation** with **teacher-student** techniques to effectively deploy on existing real-time infrastructure

**User Personalisation** (TensorFlow, PySpark, GCP)

- Personalised content delivery by developing **LSTMs** for user-level behaviours to cluster the customer base with **t-SNE**

**UNIVERSITY OF TORONTO** — Toronto, CA
Research Assistant — *Sept 2018 - Apr 2019*
**BERT for Biomedical Text** (PyTorch)

- Achieved SOTA results for biomedical **named-entity recognition** (NER) through **fine-tuning** of pre-trained **BERT** models
- Leveraged **transfer learning** and **multi-tasked learning** techniques to maximise performance across small corpora

**ANALOG DEVICES INC.** — Toronto, CA
Data Scientist (Intern) — *May 2017 - Aug 2018*
**Power Optimisation for PPG Heart Rate Sensors** (Embedded C, MATLAB)

- Designed an optimisation algorithm that reduced power consumption by 50% while maintaining sensor performance, **published at ICASPP 2020**

## EDUCATION

**UNIVERSITY COLLEGE LONDON**
**MSc Computational Statistics & Machine Learning**
Graduated with Distinction

**UNIVERSITY OF TORONTO**
**BASc Engineering Science (Robotics Specialisation)**
Graduated Honours

## RELEVANT STUDIES

Approximate Inference (PhD)
Unsupervised Learning (PhD)
RKHS's in Machine Learning (PhD)
Convex Optimisation (PhD)
Statistical Learning Theory (MSc)
Supervised Learning (MSc)
Computer Vision (MSc)

## PUBLICATIONS

Generalised Variational Inference for Gaussian Processes (Msc Thesis)

Biomedical Named-Entity-Recognition (BASc Thesis)

Power Optimization with Photoplethysmography Signal Quality Classification (ICASSP 2020)

Robust Beat-To-Beat Detection Algorithm for Pulse Rate Variability (ICASSP 2018)

3D-Printed Neuronavigation Headset for Therapeutic Brain Stimulation (JNE 2018)

## SOFTWARE SKILLS

**Fluent**
Python, Git, LaTeX

**Proficient**
Airflow, SQL, Docker, PySpark, GCP

**Familiar**
Kubeflow, MATLAB, C, Assembly

## PERSONAL INFO

Canadian Citizen
Native English Proficiency