

Sallam Saka

sallamsaka.com | sallam.saka@mail.utoronto.ca | (226) 975-3592

SKILLS

- **Technologies:** Python, R, SQL, Bash, Git, Linux, AWS, Tableau, Microsoft Office (Excel, PowerPoint)
- **Libraries:** PyTorch, TensorFlow, Keras, NumPy, Pandas, Xarray, Scikit-learn, SciPy, Matplotlib, re
- **Certifications:** SQL, Databases, AWS (Cloud, AI) Practitioner, Linux and Bash Scripting, Excel, Tableau

EDUCATION

University of Toronto

Expected 2028 (2027 + Co-op term)

HBSc, Biological Physics Specialist; Minors in Statistics and Mathematics

cGPA: **3.98/4**

- Enrolled in the Arts and Science Internship Program (Co-op Program)
- **Coursework:** Time Series Analysis (Grad-level), Probabilistic ML (Grad-level), Theory of Statistical Practice
 - Probabilistic modeling (graphical models, HMMs, MCMC, variational inference, Gaussian processes) and hands-on implementation of HMC, SVI, EM, and classical ML (PCA, GDA, k-means)

Deep Learning Specialization Certification - DeepLearning.AI

- Applied transfer learning, YOLO detection, and neural style transfer to solve advanced vision tasks
- Built and trained RNNs, LSTMs, and GRUs for text generation, music synthesis, and speech detection

EXPERIENCE

Machine Learning Researcher – Sunnybrook Research Institute

May 2026 – Present

- Built a deep learning system detecting sleep apnea from a **Muse S headband** instead of in-lab equipment
- Trained on **400+ overnight recordings**, each 10,000+ timesteps, against clinical polysomnography
- Preprocessed raw **EEG, PPG, and accelerometry** into model-ready sequences with a Python pipeline
- Implemented a CNN encoder and 4-layer bidirectional Mamba sequence model in **PyTorch**
- Ran hundreds of hyperparameter sweeps and ablations on **H100 GPUs** via SLURM across Linux HPC clusters
- Built **experiment tracking** capturing configs, metrics, and diagnostics across hundreds of runs

Research Assistant – Department of Physics, University of Toronto

May 2025 – August 2025

- Compared satellite and ground-based atmospheric ozone measurements to assess inter-instrument consistency
- Built a **Python** pipeline (**Pandas**, **NumPy**, **Xarray**) that quality-filtered **1.2M+** ozone profile records
- Co-located **14K+** coincident measurement pairs by haversine distance within 500 km and 8 hours
- Established satellite–ground ozone column agreement (**r=0.88–0.98**) with no significant 20-year drift
- Visualized results with **Matplotlib** by creating **310+** diverse plots (time series, climatology, scatter, line)
- Maintained reproducible code on GitHub and conveyed findings via a presentation and final report

PROJECTS

Microscopy Cell Counting with Density Maps (Web App, GitHub)

- Built a U-Net style model in **PyTorch** to predict cell density maps and counts from microscopy images
- Generated training labels from cell instance masks; doubled training data via contrast-inverted images
- Trained on 256×256 patches and reconstructed full-image predictions by stitching patches at inference
- Tried multiple loss formulations (pixel-level + distribution shape + count MAE); stabilized training with AdamW, gradient clipping, and ReduceLROnPlateau, selecting the best checkpoint by full-image MAE
- Achieved validation count MAE of **5.21** (full images) and **3.02** (256×256 patches)

Sensory Adaptation in Changing Environments (Research Project)

- Built a from-scratch **Python** simulation to model a neuron's adaptation to a changing environment
- Implemented online **Bayesian inference** so the neuron tracks its changing environment from noisy stimuli
- Optimized the neuron's encoding so stimuli could be reconstructed from its responses with minimal error
- Compared different rules for when the neuron adapts its encoding against ideal and non-adaptive baselines
- Evaluated each strategy on multiple accuracy measures over thousands of timesteps and 20 random seeds