Amy X. Lu

http://amyxlu.github.io
github.com/amyxlu

lu.amy326@gmail.com
+1-647-968-6833 or +1-617-415-8777
linkedin.com/in/amyxlu

Education

- **University of Waterloo** — Waterloo, Canada
  Bachelors of Science, Honours Science, Bioinformatics Option — *2014 – 2019*
- **University of Toronto, Vector Institute** — Toronto, Canada
  Masters, Computer Science — *2019 – Present*

Experience

- **Harvard Medical School/Boston Children's Hospital** — Boston, USA
  Research Intern — Deep Learning in Clinical Genomics, Sliz Lab — *2018 – Present*
  - **Machine learning:** Building machine learning models to predict disease phenotype from exome variants, with an emphasis on capturing genotype-phenotype non-linearities arising from epistasis.
  - **Dimensionality reduction:** Tackling the problem of high features-to-samples ratio in clinical genomics using both machine learning and bioinformatics techniques
  - **Identifying pathogenic variants:** Analyzing important model features to identify potentially disease-associated genomic variants.
  - **Designing computer programs:** Creating a simple command line interface enabling users to conduct post-GATK filtering and train models on variant data.
- **University of Waterloo** — Waterloo, Canada
  Undergraduate Thesis Student — Deep Learning in Regulatory Genomics, Doxey Lab — *2017 – 2018*
  - **Convolutional neural networks:** Implemented a convolutional neural network which recognizes genomic enhancers regulating femur growth.
  - **Interpreting the "black box":** Addressed the "black box" problem in neural networks by traversing all 8-mer permutations, examining their activations after the first convolutional layer, and reconstructing position-weighted matrices of genomic motifs recognized by the network.
  - **91% accuracy:** Achieved 91% model accuracy (97% AUROC) through hyperparameter modifications.
  - **Bioinformatic pipelines:** Used various bioinformatics tools to statistically match all reconstructed motifs recognized by the network to a database of experimentally-confirmed enhancers.
  - **Self-directed learning:** Self-learned deep learning and Python as a biology student.
- **École polytechnique fédérale de Lausanne** — Lausanne, Switzerland
  Research Intern — Computational Biology, Laboratory for Biomolecular Modelling — *2017*
  - **Molecular dynamics:** Contributed three potential findings on enzyme-membrane mechanisms of bacterial "superbugs" using computational simulations based on molecular dynamics (MD) principles.
  - **Automating workflows:** Increased workflow efficiency by writing R, tcl/tk, and Csh scripts to conduct analyses in parallel.
- **University of Toronto** — Toronto, Canada
  Research Intern — Data Visualization in Pharmacoepidemiology, Cadarette Lab — *2015 – 2017*
  - **Publications:** Took initiative to redesign manual workflow and exceeded expectations by contributing two co-authored publications, one first-author poster, and figures for an international conference within two months.
  - **Data visualization:** Visualized and wrangled data using R and SQL; coded interactive HTML widgets using Shiny.
- **University of Toronto** — Toronto, Canada
  Ophthalmology Research Assistant — Statistical Analysis of Surgical Efficacy, Dr. Eric Tam — *2015 – 2016*
  - **Statistical clinical analysis:** Statistical analyses on the efficacy of new prophylactic procedures in femtosecond-laser assisted cataract surgery. First-author poster submitted to AAO 2016.

Awards

- **Canada Graduate Scholarships-Master's (CGS-M) Award**: NSERC federal research grant. Valued at $17,500.
- **Scholarship of Excellence in Research**: One of 13 selected by EPFL (global rank: 12). Valued at CHF 4500.
- **University of Waterloo**: Various entrance awards totaling $6000.
- **Associate of The Royal Conservatory (ARCT)**: Performer's Diploma in Piano (highest academic standing offered).

Publications

- **Published:** Ban J, Tadrous M, **Lu AX**, Cicinelli EA, Cadarette SM. Diffusion of indirect comparison meta-analytic methods to study drugs: a systematic review and co-authorship network analysis. BMJ Open.
- **Under revision:** Consiglio GP, Maclure M, **Lu AX**, Cicinelli EA, Ban JK, McCarthy L, Cadarette SM. Guidance documents for the application of Self-controlled Crossover Observational PharmacoEpidemiology (SCOPE) designs are needed: systematic review and animated co-authorship networks. Pharmcoepi Drug Saf.
- **Under preparation: Lu AX**, Rockowitz S, Poduri A, Sliz P. From data to precision medicine: predictive machine learning models to uncover disease-associated variants

Posters

- **Presented: Lu AX**, Rockowitz S, Poduri A, Sliz P. From data to precision medicine: predictive machine learning models to uncover disease-associated variants. *Talk and poster presentation*, Harvard Medical School BCMP Retreat 2019, Whitefield, New Hampshire, Oct 2018.
- **Presented: Lu AX**, Consiglio GP, Cadarette SM. Dynamic Visualization in Co-Authorship Network Analysis. *Poster presentation*, Undergraduate Summer Research Program, Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, August 2016.
- **Presented:** McIlroy-Young R, **Lu AX**, Guenther N, Olarnyk A. Did the Arab Spring Impact the Academic Network of Tunisia Between 2010 and 2015?. *Poster presentation*, Faculty of Environment, University of Waterloo, April 2015.
- **Submitted: Lu AX**, Tam ES. Effect of prophylactic brimonidine on subconjunctival hemorrhage in laser-assisted cataract surgery. *Poster abstract*, American Academy of Ophthalmology 2017.

Service and Activities

- **SV.AI:** Core Team of SV.AI, a non-profit connecting patients of rare genomic diseases to the medical/AI research community and industry partners through collaborative research initiatives. *Ongoing.*
- **Residence Don, Velocity Incubator:** Leader for the residence program of Canada's most productive start-up incubator; fostered close relations with diverse students to support conduct and mental health crises
- **Tosamaganga Hospital, Tanzania:** Supported operations at a rural Tanzanian hospital and shadowed surgical procedures. Expenses were covered by scholarships, fundraising, and part-time tutoring.
- **Teaching:** Taught weekly tutorial lectures for a 200-level genetics course.

Hackathons and Projects

- **NLP-seq:** Using n-grams, a feature extraction technique in natural language processing (NLP), for supervised learning of function prediction using Protein Data Bank data.
- **Siamese networks:** Training Siamese networks, conventionally used to differentiate if two images are of the same face, on enhancers regulating the proximal and distal femur, in order to quantify their differences.
- **Chatsense:** Mobile chat app using a sentiment analysis machine learning API on voice messages and mapping dominant emotions to emojis, designed for the autistic community. *Team project. Winner at YHacks at Yale.*
- **Alice:** Web login page using a trained facial recognition model to login to online banking without passwords, designed for dementia patients. *Team project. Winner at ConUHacks III.*

Programming Skills

- **Machine learning**: Keras, TensorFlow, scikit-learn, NumPy
- **Data visualization**: Matplotlib, ggplot2, igraph
- **Languages, tools**: Python, R, Bash, C, Vim, Jupyter, SQL
- **Computational biology**: GROMACS, Visual Molecular Dynamics, HMMER, ClustalW, I-TASSER