JONATHAN KING

Education

Carnegie Mellon - University of Pittsburgh Joint PhD Program in Computational Biology

University of California, Berkeley

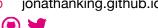
B.A. Computer Science, B.S. Bioengineering

May 2017

Expected Oct 2023

Skills and Coursework

510.856.7311 tel jonathanking.jek@gmail.com email jonathanking.github.io



- · Python, PyTorch/Lightning, Tensorflow, pandas, Docker, PyMol, Slurm HPC, NumPy, ProDy, Spark/Dask, C, Java, SQL • Discrete Math, Probability,
- · Deep Learning, Cloud Computing, **Neural Machine Translation**
- · Molecular, Systems, and Structural Biology, Biochemistry
- Linear Algebra, Operating Systems, Databases, Algorithms

Aug 2017 - Present

Research Experience

Graduate Student Researcher

Carnegie Mellon - University of Pittsburgh (David Koes Lab)

- Developed SidechainNet A, an open-source Python package, dataset, and suite of tools to fast-track machine learning research in protein science and lower the barrier to entry for the community.
- Spearheaded novel deep learning (DL) techniques to generate all-atom protein structure predictions with recurrent and attention-based sequence models (RNNs, Transformers), facilitating drug discovery.
- Devised DL methods that incorporate force fields (OpenMM, AMBER) to improve models like AlphaFold2.
- Explored analogous methods in RNA structure prediction, sequence generation, and property prediction.
- Recognized as a skilled communicator and teacher, able to break down complex ideas for diverse audiences.

May 2020 - Aug 2020 Research Intern

Google Brain

- Developed large-scale DL models to accurately predict complex phenotypes (e.g., diseases like COPD) from noisy, multi-modal biomedical data, reducing the costs of physician-labeled data.
- Investigated models' ability to identify both known and novel genetic variants linked with COPD via GWAS.
- · Collaborated with a team of expert researchers and engineers, contributing to a high-impact project that has the potential to revolutionize personalized medicine.

Research Assistant Dec 2016 - Aug 2017

University of California, San Francisco

- Pioneered RNN methods for cardiac arrest prediction with AUROC=0.85, enhancing preventive care.
- Fine-tuned hospital alert systems by examining varying prediction timeframes, reducing alarm fatigue.

Bioinformatics Intern May 2015 - Dec 2016

Plexxikon Inc., Berkeley, CA

- Formulated novel string-based algorithms to detect structural variants in acute myeloid leukemia patients.
- Deployed program on Illumina's BaseSpace cloud-computing platform with Amazon Web Services.
- · Optimized algorithms' specificity and reporting methods, outperforming existing software on target genes.

Research Assistant Apr 2014 - Sep 2014

University of California, Berkeley

- Examined epigenetic role of DNA methylation in Arabidopsis thaliana via PCR and computational analysis.
- Identified phenotypic responses of 40 previously uncharacterized chromatin remodeler protein mutants.

Other Projects

instaGAN: De Novo Food Blogging with Generative Models

Apr 2019

Combined Generative Adversarial Networks and RNNs to generate and caption Instagram-style food photos.

RNA Secondary Structure Prediction via Neural Machine Translation

Apr 2018

- Created RNN methods to predict RNA structure and terminator strength from sequence.
- Utilized multi-task learning to develop a shared latent space model, improving AUROC by ~0.1.

Leadership

Summer Research Mentor

2018, 2019, 2022, 2023

University of Pittsburgh, TECBio Research Experience for Undergraduates

- Mentored 4 students with graduate-level research projects in machine learning & drug discovery.
- · Advised from start to finish, culminating in presentations at university and minority student research symposia.

Leadership cont.

Summer Research Ethics Forum & Journal Club Mentor

2018, 2019, 2022

University of Pittsburgh, TECBio Research Experience for Undergraduates

• Guided students through an interdisciplinary forum on research ethics; nurtured scientific presentation skills.

Tutor - Computer Science, Math, Chemistry, Physics

Dec 2014 - 2019

InstaEdu.com, Online Tutoring Service

• Coached 40+ high school and college students in STEM fields, crafting weekly lesson plans.

Publications

SidechainNet: An All-Atom Protein Structure Dataset for Machine Learning.

King, J. E., Koes, D. Proteins: Structure, Function, and Bioinformatics (2020).

DREAMing of big data and scalable machine learning: Predicting kinase binding with matrix factorization. Koes, D., **King, J. E.**, Francoeur, P. G., Kowalczyk, A., Rajashekar, S., Chennubhotla, C. *Abstracts of Papers of the American Chemical Society* (2019).

Convolutional neural network scoring and minimization in the D3R 2017 community challenge. Sunseri, J., **King, J. E.**, Francoeur, P. G., Koes, D. *Journal of Comp. Aided Molecular Design* (2018).

Predict In-Hospital Code Blue Events using Monitor Alarms through Deep Learning Approach. Xiao, R., **King, J. E.**, Villaroman, A., Do, D. H., Boyle, N. G., Hu, X. *IEEE Eng. in Med. and Bio.* (2018).

Awards

Best Elevator Pitch, Best Talk, Best Poster, CPCB PhD Program Annual Meeting 2019, 2020, 20		
Natl. Inst. of Biomed. Imaging and Bioeng. T32 Training Grant		Sep 2019 - Sep 2021
Biomedical Graduate Student Association, Travel Award	Aug 2019, Nov 202	1, Mar 2022, Mar 2023
Best Talk, Northern California Computational Biology Student Symposium		Oct 2016

Presentations (Conferences and Symposia)

SidechainNet: Training Dec	p Learning Protein Structure	Models with Force Fields
----------------------------	------------------------------	---------------------------------

American Chemical Society (ACS), Computational Chemistry Division

Mar 2023

SidechainNet: An All-Atom Protein Structure Dataset for Machine Learning

ACS, Computational Chemistry Division
AlgoSB: Proteins and Machine Learning @ CIRM in Marseille, France
NeurIPS, Machine Learning for Structural Biology Workshop
Al-Driven Discovery in Biophysics, Carnegie Mellon University
PhD Intern Research Conference, Google Brain

Mar 2022
Nov 2021
Dec 2020
Jun 2020

Exploring sequence-to-sequence learning methods for end-to-end, complete protein structure prediction

ACS, Computational Chemistry Division

Canadian Chemistry Conference, Machine Learning Division

Jun 2019

University of Pittsburgh Advanced Research in Computing Symposium

Mar 2019

A Novel Algorithm for Detecting FLT3 Internal Tandem Duplications in Acute Myeloid Leukemia Patients

Northern California Computational Biology Student Symposium Oct 2016

Presentations (as Invited Speaker)

Deep Learning for Protein Structure Prediction

University of Pittsburgh, Scalable Machine Learning, Guest Lecturer

UC Berkeley, Comp Bio Skills Seminar and Berkeley Connect Comp Bio

Weights and Biases, Deep Learning Salon, YouTube

Sidley Austin LLP, Technology and Life Sciences Transactions Informational Meeting

Mar 2021 - 2023

May 2021, Apr 2022

Apr 2020

May 2023

Presentations (as Summer Research Mentor)

Screening and Simulating Potential Inhibitors for the CYP4F2 Enzyme

Developing a Latent Space Representation for Prediction of both RNA Terminator Strength and Structure

Presenters: Jackelyne Garcia Cruz, Alex Ludwig

Annual Biomedical Research Conference for Minority Students

Nov 2019
Summer Undergraduate Research Symposium, Duquesne University

Jul 2018, Jul 2019