Sonny Arora, Ph.D.

GitHub: https://github.com/sonnyarora

EXPERIENCE SUMMARY: Data Scientist and Mathematician with 5+ years' experience applying machine learning models and analysis to large scale data in biochemistry.

WORK EXPERIENCE

Pennsylvania State University **July 2019 – June 2022**
Computational Scientist *University Park, PA*

- Applied machine learning models (LSTM/convolutional neural networks, SVMs, linear models) to the problem of protein/DNA binding to understand what model architectures/training schemes/loss functions improve performance.
- SeqView: Created web-frontend using Django/Javascript(React)/MySQL for displaying terabytes of genomic data (bigwigs); website used by lab and collaborators.

Skills: Python, Tensorflow, Django, React, MySQL, Docker, Docker-compose.

Emory University **August 2018 – September 2019**
Visiting Assistant Professor *Atlanta, Georgia*

- Taught courses in Integral/Multivariable Calculus; Used examples from probability/statistics in lectures/homework.

Pennsylvania State University **August 2012 - August 2018**
Graduate Teaching Assistant *University Park, PA*

- Created an algorithm and proved its effectiveness for constructing curves for cryptography; Implementation using math programming languages MAGMA and Sage (Python). Published work in journal.
- With coauthors, proved a statement about probability distributions of dimension 3 objects (Twisting Sato-Tate Conjecture for a specific curve); published work in journal.

Skills: MAGMA, Python, C++.

Mevio. Inc. **September 2011-May 2012**
Software Engineer *San Francisco, CA*

- Extended existing pipeline to generate reports from website traffic logs using Hadoop in AWS; Reports were loaded into MongoDB database for viewing in data analytics dashboard saving hours of work a week for business analysts.

Skills: Javascript, PHP, Hadoop, MySQL, AWS.

EDUCATION

Pennsylvania State University **May, 2018**
Ph.D. in Mathematics *University Park, PA*

University of California - Davis **May, 2011**
B.S. in Computer Science, B.S. in Mathematics *Davis, CA*

SKILLS

Python; R; Bash; Javascript; C++; Pytorch; Huggingface Transformers library; TensorFlow; ScikitLearn; Numpy; Pandas; Django; matplotlib; seaborn; MySQL; Docker; Docker compose; Git; GitHub; MongoDB; Hadoop; AWS; Azure

PUBLICATIONS

Arora, S., & Eisenträger, K. (2019). Constructing Picard curves with complex multiplication using the Chinese remainder theorem. *The Open Book Series*, *2*(1), 21-36.

Arora, S., Cantoral-Farfán, V., Landesman, A., Lombardo, D., & Morrow, J. S. (2018). The twisting Sato–Tate group of the curve y^2= x^8-14x^4+ 1. *Mathematische Zeitschrift*, *290*, 991-1022.

PUBLICATIONS IN PREPARATION

Arora, S., Yang, J., Akiyama, T., James, D., Morrissey, A., Blanda, T., Badjatia, N., Lai, WK., Ko, M., Pugh, BF., Mahony, S., Joint sequence and chromatin neural networks characterize the differential abilities of Forkhead transcription factors to engage inaccessible chromatin.