Kyle Alford

(313)-409-1230 | kla2122@columbia.edu
www.linkedin.com/in/klalford/

EDUCATION

Columbia University in the City of New York — **New York, NY**
Bachelor of Science in Computer Science, Statistics — Expected 05/2023

- **GPA**: 3.8 / 4.0
- **Relevant Coursework**: Artificial Intelligence, Computer Vision, Bayesian Statistics, Statistical Inference, Natural Language Processing, Machine Learning for Functional Genomics, Data Structures & Algorithms, Statistics & Probability

WORK EXPERIENCE

Microsoft — **Redmond, WA**
Software Engineering Intern — 06/2022 – Present

- Azure Stack Hub is an edge compute cloud offering of Azure that provides disconnected cloud computing services
- Designed and implemented 15 metrics such as latency, error rate, number of retries, etc., using the ETW framework, C# and .NET, to monitor virtual machine health
- Created dashboard visualizations for these metrics using Azure Grafana, a visualization tool used internally by engineers to debug customer VM errors
- Presented the dashboards to entire Azure Stack organization, allowing engineers to root cause VM errors before they reach customers

Rajpurkar Lab @ Harvard Medical School Biomedical Informatics — **Boston, MA**
Deep Learning Researcher — 12/2021 – Present

- Monitored and improved an existing convolutional neural network model that diagnoses lung diseases based on X-ray images
- Redeployed the model in two understaffed hospitals in Nam Dinh, Vietnam which improved accuracy and time-efficiency in treating 20,000 patients
- Analyzed model outputs to discover out-of-distribution X-ray subpopulations to reduce disagreement between AI and radiologists

Shah Lab @ Memorial Sloan Kettering Cancer Center — **New York, NY**
Software Engineering Intern — 05/2021 – 09/2021

- Reconfigured lab software, a Hidden Markov Model that segments tumor cells into 12 states of copy number using single cell genomics data, from R into Python (PyTorch, Pyro)
- Reduced the run-time from 12 hours to 3 hours using parallelization and replaced complicated hyperparameters with most likely values using Stochastic Variational Inference for improved UX
- Presented improved software to several computational biology labs and clinician teams that currently use it

Kumar Lab @ University of Miami Public Health — **Coral Gables, FL**
Biostatistics Intern — 09/2019 – 02/2021

- Investigated the effects of volatile organic compounds (VOCs) and other air pollutants on respiratory health in metropolitan cities like Miami, FL
- Collected bacterial samples from patient homes, identified them with online databases, and used hypothesis tests to determine which bacteria significantly associate with asthma
- Published a meta-analysis to corroborate our findings and performed population statistical analysis using R

PROJECTS

PyTorch Diagnose Blood Cancer from ATAC-seq data using Autoencoder and CNN

- Used data from Cardiovascular Disease Knowledge Portal (CVDKP) which contains chromatin accessibility and phenotype-level info for several thousand patients
- Created linear autoencoder to reduce size of data and discover linear combinations of features that varied the most, then mapped new features to identifiable regions of genome for interpretability
- Designed CNN for binary classification task (cancerous/non-cancerous)

Tensorflow Disease Classification on X-rays using Transformer

- Used Chest X-ray images from CheXpert dataset of patients that are either healthy or have varying stages of lung diseases
- Trained a Visual Transformer (ViT) model and implemented DeepAUC loss from scratch to achieve 90% accuracy

LEADERSHIP

Columbia Data Science Society — **New York, NY**
Club President — 09/2020 – Present

- Organized Hackathon which brings 400+ virtual participants each year and liaise between tech company sponsors for funding
- Created career development and enrichment opportunities for the 25 club members
- Received 2021-22 Zvi Galil award for Best Engineering Club of the Year

Stanford AI4ALL — **San Francisco, CA**
Portfolio Project Mentor — 05/2022 – Present

- Mentored a group of 8 high schoolers in developing their interests/skill in Machine Learning

- Advised over their projects in areas of Computer Vision and NLP, then prepared them for presenting their work

SKILLS, ACTIVITIES & INTERESTS

- **Languages:** Fluent in English; Proficient in Spanish, Hindi, Urdu
- **Technical Skills:** Python, R, Java, C/C++, SQL, .NET, C#, Bash, JavaScript; Git, Docker
- **Activities**: ICLR 2022 Volunteer, National Society of Black Engineers, Columbia Organization of Rising Entrepreneurs
- **Awards:** IBM Accelerate SWE Fellow, Sol Spiegelman Pre-Medical Excellence Scholarship, Columbia University Dean's List
- **Interests:** Playing Drums, Cricket, Basketball, Swimming, Chess, True Crime, Learning Languages, Puzzles