

Ulzee An

PhD Student

Computer Science Department

University of California, Los Angeles

[ulzee@cs.ucla.edu](mailto:ulzee@cs.ucla.edu)

[0000-0001-8066-7291](tel:0000-0001-8066-7291)

<https://www.linkedin.com/in/ulzee/>

## Preprints

- **An, U.**, Jeong, M., Lee, S., Gorla, A., Chiang, J.N. and Border, R. (2024). “BrainSplat: Bootstrapping highly scalable and expressive brain MRI embeddings from foundation models.”
- **An, U.**, Gorla, A., Jeong, M. and Sankararaman, S. (2024). “Gaussian process-based neural topic model for series-specific topic dynamics in time series datasets.”
- Jew, B., Korb, A., Lou, P., Chiang, J.N., **An, U.**, Sahai, A., Halperin, E. and Eskin, E. (2020). “Expanding COVID-19 symptom screening to retail, restaurants, and schools by preserving privacy using relaxed digital signatures.”
- Chiang, J.N., **An, U.**, Kordi, M., Jew, B., Kravit, C., Dunne, W.J., Perez, R., Parikh, N.R., Weil, D., Azar, R.F. and Cherry, R. (2020). “Projecting hospital resource utilization during a surge using parametric bootstrapping.”

## Publications

- **An, U.**, Lee, S., Gorla, A., Jeong, M. and Sankararaman, S. (2024). “DK-BEHRT: Teaching Language Models International Classification of Disease (ICD) Codes using Known Disease Descriptions” *AAAI workshop on AI for Medicine and Healthcare*
- Avram, O., Durmus, B., Rakocz, N., Corradetti, G., **An, U.**, Nittala, M. G., Terway, P., Rudas, A., Chen, Z.J., Wakatsuki, Y., Hirabayashi, K., ..., Halperin, E. (2024). “Accurate prediction of disease-risk factors from volumetric medical scans by a deep vision model pre-trained with 2D scans.” *Nature Biomedical Engineering*
- **An, U.**, Pazokitoroudi, A., Alvarez M., Hunag, L., Bacanu, S., Schork, J. A., Kendler, K., Pajukanta, P., Flint, J., Zaitlen, N., Cai, N., Dahl, A., Sankararaman S. (2023). “Deep Learning-based Phenotype Imputation on Population-scale Biobank Data Increases Genetic Discoveries.” *Nature Genetics*
- Dahl, A., Thompson, M., **An, U.**, Krebs, M., Appadurai, V., Border, R., Bacanu, S.-A., Werge, T., Flint, J., Schork, A. J., Sankararaman, S., Kendler, K., & Cai, N. (2023). “Phenotype integration improves power and preserves specificity in biobank-based genetic studies of MDD.” *Nature Genetics*
- Chiang, J. N., Corradetti, G., Nittala, M. G., Corvi, F., Rakocz, N., Rudas, A., Durmus, B., **An, U.**, Sankararaman, S., Chiu, A., Halperin, E., & Sadda, S. R. (2023). “Automated Identification of Incomplete and Complete Retinal Epithelial Pigment and Outer Retinal Atrophy Using Machine Learning.” *Ophthalmology Retina*
- **An, U.**, Cai, N., Dahl, A., & Sankararaman, S. (2022). “AutoComplete: Deep Learning-Based Phenotype Imputation for Large-Scale Biomedical Data.” *Research in Computational Molecular Biology (Recomb)*
- **An, U.**, Shenhav, L., Olson, C. A., Hsiao, E. Y., Halperin, E., & Sankararaman, S. (2022). “STENSL: Microbial Source Tracking with ENvironment SeLection.” *mSystems*
- **An, U.**, Bhardwaj, A., Shameer, K., & Subramanian, L. (2021). “High Precision Mammography Lesion Identification From Imprecise Medical Annotations.” *Frontiers in Big Data*
- Goodman-Meza, D., Rudas, A., Chiang, J. N., Adamson, P. C., Ebinger, J., Sun, N., Botting, P., Fulcher, J. A., Saab, F. G., Brook, R., Eskin, E., **An, U.**, ..., and Manuel, V. (2020). “A machine learning algorithm to increase COVID-19 inpatient diagnostic capacity.” *PLOS ONE*
- Iyer, S. R., **An, U.**, & Subramanian, L. (2020). “Forecasting Sparse Traffic Congestion Patterns Using Message-Passing RNNs.” *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*

## Education

---

2019-current	PhD Student <i>Computer Science</i> <i>Advised jointly by Sriram Sankararaman</i> <i>University of California, Los Angeles</i>
2017-2019	Master of Science <i>Computer Science</i> <i>Advised by Lakshmi Subramanian</i> <i>Courant Institute of Mathematical Sciences, New York University</i>
2012-2016	Bachelor of Science <i>Computer Science</i> <i>University of Illinois, Urbana Champaign</i>

## Professional Experience

---

Research Scientist Intern   <b>Uber</b>   San Francisco, CA	June 2022 – Sep 2022
<ul style="list-style-type: none"><li>Developed probabilistic and deep learning models to learn an embedding space for all cities serviced by Uber, using hourly real-time metrics of service demand and efficiency.</li></ul>	
Cofounder   <b>PiraShield</b>   Boston, MA	June 2016 – July 2017
<ul style="list-style-type: none"><li>Boston MassChallenge 2016 Finalist, awarded \$100,000 in AWS credits</li><li>Architected a distributed web crawler to index 10k+ domains serving pirated media (over 2 million pages) protected behind anti-crawling firewalls</li></ul>	
Software Engineering Intern   <b>Rithmio</b>   Chicago, IL	May 2015 – Sep 2015
<ul style="list-style-type: none"><li>Developed a C language SDK for a proprietary motion recognition algorithm using 3-axis gyroscope sensor data. Deployed the algorithm to the Apple iPhone, Apple Watch and Pebble Smartwatch.</li></ul>	

## Presentations

---

- An, U.**, Fu, J., Eskin, E., Sankararaman, S. (2024). “ImputeMLP: Towards an efficient and scalable genotype imputation method using deep learning.” Poster presented at the 2024 annual meeting of the American Society of Human Genetics (ASHG).
- An, U.**, Lee, S., Gorla, A., Jeong, M. and Sankararaman, S. (2024). “Medical BERT training with known disease representations from biomedical language models” Poster presented at the UCLA-Amazon Science Hub 2024 Showcase.
- An, U.**, Pazokitoroudi, A., Alvarez, M., Huang, L., Bacanu, S., Schork, A. J., Kendler, K., Pajukanta, P., Flint, J., Zaitlen, N., Cai, N., Dahl, A., & Sankararaman, S. (2022). “Deep Learning-based Phenotype Imputation on Population-scale Biobank Data Increases Genetic Discoveries.” Poster presented at the 2022 annual meeting of the American Society of Human Genetics (ASHG).
- An, U.**, Cai, N., Dahl, A., & Sankararaman, S. (2022). “AutoComplete: Deep Learning-Based Phenotype Imputation for Large-Scale Biomedical Data.” Talk at the 2022 annual meeting of Research in Computational Molecular Biology (Recomb).

## Other Contributions

---

- Iyer, S. R., Balashankar, A., Aeberhard, W. H., Bhattacharyya, S., Rusconi, G., Jose, L., Soans, N., Sudarshan, A., Pande, R., & Subramanian, L. (2022). *Modeling fine-grained spatio-temporal pollution maps with low-cost sensors*. Npj Climate and Atmospheric Science, 5(1), 76. doi:10.1038/s41612-022-00293-z  
Acknowledged for implementing and testing related works.
- StopCovid19Together.com. Technical lead (development and deployment of public facing website, backend, and algorithms). A major interdisciplinary undertaking between UCLA Health and UCLA Computational Medicine to monitor and project the spread of Covid-19 during its initial spread.

## Awards

---

### Amazon Fellow

June 2023

Fellowship awarded through the Science Hub for Humanity and Artificial Intelligence (a partnership between UCLA and Amazon).

### National Library of Medicine Data Science-Precision Health Fellowship

July 2023

(declined - conflict with Amazon Fellowship) Fellowship awarded by Institute for Precision Health and Computation Medicine at UCLA, funded by the National Library of Medicine.

## Teaching Positions

---

2023	Introduction to Machine Learning (Undergraduate) <i>Teaching Assistant</i> <i>Computer Science Department, University of California Los Angeles (UCLA)</i>
2022	Introduction to Machine Learning (Undergraduate) <i>Teaching Assistant</i> <i>Computer Science Department, University of California Los Angeles (UCLA)</i>
2018	Introduction to Data Science (Graduate) <i>Teaching Assistant</i> <i>Center for Data Science, New York University (NYU)</i>