DATA SCIENCE INSTITUTE

Data Science Community Hour

Weekly in-depth technical discussions, presentations, tutorials, and meet-ups from the community of data science researchers and educators. Students, staff, researchers, and faculty are welcome!

Thursday April 1st, 4:30–5:30 pm via Zoom (click to join)

Interview with Dr. Moumita Bhattacharya @Netflix

Moumita Bhattacharya, Ph.D., is a Senior Research Scientist at Netflix where she develops machine learning models for search and recommendations algorithms and is a member of the core AI/ML research team. Moumita recently joined Netflix. Prior to Netflix, she was a Senior Applied Scientist at Etsy, a two-sided marketplace for buyers and sellers. At Etsy, Moumita was the tech lead of a team that develops recommendation systems to show relevant products to Etsy users. Many of her developed approaches at Etsy improved conversion rates and gross merchandise sales of the company. Moumita received the Ph.D. degree in computer science from the University of Delaware, where she was a member of the Computational Biomedicine and Machine Learning Lab, led by her thesis advisor Prof. Hagit Shatkay. During her Ph.D., Moumita focused on developing machine learning models for disease prediction and patient risk stratification, including risk prediction for heart disease and chronic kidney disease, in a collaboration with cardiologists from Johns Hopkins University and nephrologists from the largest hospital in Delaware. Prior to grad school, Moumita worked as a software engineer for three years in the financial services industry.

"How to Approach Your First Data Science Project" — Mauricio Ferrato, CIS, UD.
This talk was prepared with the intention of helping researchers who are new to the field to start their first data science project. The speaker will talk about their experience on a few data science projects, discussing the do's and don'ts of approaching these types of projects. The talk is built around the machine learning workflow where the best practices for data preprocessing, feature selection, classification/regression, hyperparameter tuning, and validation will be presented. The presenter will also talk about the state-of-the-art tools and techniques available and which to use depending on your data.

Mauricio Ferrato is a third year Ph.D. student in the Department of Computer and Information Sciences at the University of Delaware. He is part of the Computational Research and Programming Lab (CRPL) led by Dr. Sunita Chandrasekaran. His research interests are in high performance computing and machine learning. Mauricio has experience accelerating scientific applications using directive-based parallel programming models, as well as building and scaling machine learning/deep learning frameworks to run in large computer systems. Currently, Mauricio is working in collaboration with the Nemours A.I DuPont Hospital for Children building robust machine learning and deep learning models to predict disease onsets and outcomes for pediatric patients with rare diseases.