THE UNIVERSITY OF GEORGIA
DEPARTMENT OF STATISTICS
Colloquium Series

4:00 PM, Thursday, November 3, 2022, via Zoom

https://zoom.us/j/93480771127?pwd=cFpzVDlkWWlBM2ZPaXdIZkF5d2c4dz09

Meeting ID: 934 8077 1127 Passcode: 371551

Dr. Claire Donnat

Department of Statistics, University of Chicago

Leveraging Horizontal and Vertical Network Structure

The recent years have witnessed a surge in the amount of available structured data, typically modelled as a network capturing the relationships between different entities. This structure can either hold "horizontally" across features, or "vertically" across observations, and can be leveraged to considerably improve estimation—two aspects that we propose exploring throughout this talk. In the first part of the talk, we will focus on using network structure as a regulariser. We will consider a novel $\ell 1 + \ell 2$-penalty, which we refer to as the Generalized Elastic Net, for regression problems where the feature vectors are indexed by vertices of a given graph and the true signal is believed to be smooth or piecewise constant with respect to this graph. Under the assumption of correlated Gaussian design, we will derive upper bounds for the prediction and estimation errors, which are graph-dependent and consist of a parametric rate for the unpenalized portion of the regression vector and another term that depends on our network alignment assumption. In the second part of the talk, we will focus on the case where observations are assumed to lie on a graph. This setting has been more thoroughly explored recently in the Machine Learning community through the use of Graph Neural Networks (GNNs)— a generalization of the Deep Neural Network machinery to the graph setting. Yet, despite the multiplication of GNN methods across tasks and applications, the impact of their different components on their performance remains ill-understood. In particular, I will describe some recent work that attempts to understand the effect of the convolution operator used to aggregate information over entire neighborhoods on the geometry of the GNN embedding space.

About the Speaker

Dr. Claire Donnat is an Assistant Professor in the Department of Statistics of University of Chicago. Prior to coming to Chicago, she graduated from Ecole Polytechnique with an M.Sc. in Applied Mathematics in 2015, and with a PhD in Statistics from Stanford in 2020, where she was advised by Prof. Susan Holmes. Her research is at the intersection between Statistics and Machine Learning, and has focused on the statistical analysis of graphs and networks. She is especially interested in applying these methods to the analysis of biomedical data, and in particular, to neuroscience and brain connectomics.