DBDS Workshop in Biostatistics

https://stanford.zoom.us/j/93911828451?pwd=UVNISlBJMUhwdzI1dGJVSG5PYlVmUT09&from=addon

(In Person & Virtual access Password: 210310)

DATE:	April 14, 2022
TIME:	1:30-2:50pm
TITLE:	Machine Learning-Based Protein Engineering
SPEAKER:	**Jennifer Listgarten** Professor, UC Berkeley Electrical Engineering and Computer Science Department and Center for Computational Biology.

Abstract:

Data-driven design is making headway into a number of application areas, including protein, small-molecule, and materials engineering. The design goal is to construct an object with desired properties, such as a protein that binds to a target more tightly than previously observed. To that end, costly experimental measurements are being replaced with calls to a high-capacity regression model trained on labeled data, which can be leveraged in an in silico search for promising design candidates. The aim then is to discover designs that are better than the best design in the observed data. This goal puts machine-learning based design in a much more difficult spot than traditional applications of predictive modelling, since successful design requires, by definition, some degree of extrapolation---a pushing of the predictive models to its unknown limits, in parts of the design space that are a priori unknown. In this talk, I will discuss how we think of this problem, some of our methodological approaches to it, as well as report on our recent success in designing AAV-based gene therapy delivery libraries.