Invited sessions for ISCB42 (Lyon, 2021)
Session 5

The best of both worlds: combining deep learning and modeling

Organizers : Harald Binder : Institute for Medical Biometry and Statistics, Freiburg, DE

Chris Rackauckas : Department of Mathematics, Massachusetts Institute of Technology, USA
Pharmacometrics-Informed Deep Learning with DeepNLME

Austin Benson : Cornell University USA
Temporal and relational machine learning for biostatistical and other scientific applications

Göran Köber : Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center, University of Freiburg, Germany
Individualizing deep dynamic models for psychological resilience data

Pharmacometrics-Informed Deep Learning with DeepNLME

Chris Rackauckas[1,2], Vijay Ivaturi[2,3]

[1] Department of Mathematics, Massachusetts Institute of Technology, USA (presenting author)
[2] Pumas-AI, USA
[3] School of Pharmacy, University of Maryland, USA

Nonlinear mixed effects modeling (NLME) is commonly employed throughout the clinical pharmacometrics community in order to uncover covariate relationships and understand the personalized effects involved in drug kinetics. However, in many cases a full model of drug dynamics is unknown. Even further, common models used throughout clinical trials ignore many potentially predictive covariates as their connection to drug effects is unknown. Given the rise of machine learning, there have been calls to utilize deep learning techniques to potentially uncover these unknown relationships, but common deep learning techniques are unable to incorporate the prior information captured in known predictive models and thus are not predictive with the minimal data available. Thus the question: is it possible to bridge the gap between deep learning and nonlinear mixed effects modeling?

In this talk we will describe the DeepNLME method for performing automatic discovery of dynamical models in NLME along with discovery of covariate relationships. We will showcase how this extension of the universal differential equation framework is able to generate suggested models in a way that hypothesizes testable mechanisms, predicts the covariates of interest, and allows incorporating data in the form of images and sequences into the personalized precision dosing framework. This framework and the automated model discovery process will be showcased in the Pumas pharmaceutical modeling and simulation environment. We will end by describing how this is being combined with recent techniques from Bayesian Neural Ordinary Differential Equations in order to give probabilistic estimates to the discovered models and allow for direct uncertainty quantification. Together this demonstrates a viable path for incorporating all of the knowledge of pharmacometricians into the data-driven future.

Temporal and relational machine learning for biostatistical and other scientific applications

Austin R. Benson
Cornell University
arb@cs.cornell.edu

January 15, 2021

Abstract

Machine learning models have made it easier to reason about not only big, but also complex, data that is pervasive throughout science and engineering. In this talk, I will discuss some of our recent research on designing machine learning models for making predictions about data with complex structure, specifically temporal and relational data, for various applications, including biostatistics. A theme of the talk will be the advantages and disadvantages of deep neural network approaches to modeling such data. As part of this, I will highlight applications in time series forecasting where deep models have been particularly useful, as well as applications where deep models are unnecessary and computationally expensive. In the latter case, domain knowledge has enabled us to design algorithms that are faster and scale to larger datasets, without sacrificing accuracy. The talk will also cover some reasons why biostatistics has unique opportunities compared to other common application domains, such as social and information network analysis.

Individualizing deep dynamic models for psychological resilience data

Göran Köber[1], Shakoor Pooseh[2], Haakon Engen[3] et al.

[1] Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center, University of Freiburg, Germany (presenting author)
[2] Institute of Physics, University of Freiburg, Germany
[3] Neuroimaging Center, Johannes Gutenberg University Medical Center, Germany

Deep learning approaches can uncover complex patterns in data. In particular, variational autoencoders (VAEs) achieve this by a non-linear mapping of data into a low-dimensional latent space. Motivated by an application to psychological resilience in the Mainz Resilience Project (MARP), which features intermittent longitudinal measurements of stressors and mental health, we propose an approach for individualized, dynamic modeling in this latent space. Specifically, we utilize ordinary differential equations (ODEs) and develop a novel technique for obtaining person-specific ODE parameters even in settings with a rather small number of individuals and observations, incomplete data, and a differing number of observations per individual. This technique allows us to subsequently investigate individual reactions to stimuli, such as the mental health impact of stressors. A potentially large number of baseline characteristics can then be linked to this individual response by regularized regression, e.g., for identifying resilience factors. Thus, our new method provides a way of connecting different kinds of complex longitudinal and baseline measures via individualized, dynamic models. The promising results obtained in the exemplary resilience application indicate that our proposal for dynamic deep learning might also be more generally useful for other application domains.