

# Three Principles of Data Science: Predictability, Stability, and Computability

Bin Yu

Departments of Statistics and EECS  
University of California, Berkeley  
binyu@berkeley.edu

## ABSTRACT

In this talk, I'd like to discuss the intertwining importance and connections of three principles of data science in the title in data-driven decisions.

Making prediction as its central task and embracing computation as its core, machine learning has enabled wide-ranging data-driven successes. Prediction is a useful way to check with reality. Good prediction implicitly assumes stability between past and future. Stability (relative to data and model perturbations) is also a minimum requirement for interpretability and reproducibility of data driven results (cf. Yu, 2013). It is closely related to uncertainty assessment. Obviously, both prediction and stability principles cannot be employed without feasible computational algorithms, hence the importance of computability.

The three principles will be demonstrated in the context of two neuroscience projects and through analytical connections. In particular, the first project adds stability to predictive modeling used for reconstruction of movies from fMRI brain signals for interpretable models. The second project use predictive transfer learning that combines AlexNet, GoogleNet and VGG with single V4 neuron data for state-of-the-art prediction performance. Our results lend support, to a certain extent, to the resemblance of these CNNs to brain and at the same time provide stable pattern interpretations of neurons in the difficult primate visual cortex V4.

B. Yu (2013). Stability. *Bernoulli*, 19 (4), 1484-1500. (Invited paper for the Special Issue commemorating the 300th anniversary of the publication of Jakob Bernoulli's *Ars Conjectandi* in 1712).

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

*KDD'17, August 13-17, 2017, Halifax, NS, Canada.*

© 2017 Copyright is held by the owner/author(s).

ACM ISBN 978-1-4503-4887-4/17/08.

DOI: <http://dx.doi.org/10.1145/3097983.3105808>