

# On Big Data Learning for Small Data Problems

Yee Whye Teh

University of Oxford/DeepMind  
Oxford, UK  
y.w.teh@stats.ox.ac.uk

## ABSTRACT

Much recent progress in machine learning have been fueled by the explosive growth in the amount and diversity of data available, and the computational resources needed to crunch through the data. This begs the question of whether machine learning systems necessarily need large amounts of data to solve a task well. An exciting recent development, under the banners of meta-learning, lifelong learning, learning to learn, multitask learning etc., has been the observation that often there is heterogeneity within the data sets at hand, and in fact a large data set can be viewed more productively as many smaller data sets, each pertaining to a different task. For example, in recommender systems each user can be said to be a different task with a small associated data set, and in AI one holy grail is how to develop systems that can learn to solve new tasks quickly from small amounts of data.

In such settings, the problem is then how to “learn to learn quickly”, by making use of similarities among tasks. One perspective for how this is achievable is that exposure to lots of previous tasks allows the system to learn a rich prior knowledge about the world in which tasks are sampled from, and it is with rich world knowledge that the system is able to solve new tasks quickly. This is a very active, vibrant and diverse area of research, with many different approaches proposed recently. In this talk I will describe a view of this problem from probabilistic and deep learning perspectives, and describe a number of efforts in this direction that I have recently been involved in.

## Author Keywords

Machine learning; big data; small data; meta-learning; lifelong learning; learning to learn; multitask learning; learn to learn.

## BIOGRAPHY

Yee Whye Teh is a Professor of Statistical Machine Learning at the Department of Statistics, University of Oxford and a Research Scientist at DeepMind. He obtained his Ph.D. at the University of Toronto (working with Geoffrey Hinton), and did postdoctoral work at the University of California at Berkeley (with Michael Jordan) and National University of Singapore (as Lee Kuan Yew Postdoctoral Fellow). He was a Lecturer then a Reader at the Gatsby Computational Neuroscience Unit, UCL, and a tutorial fellow at University College Oxford, prior to his current appointment. He was programme co-chair of ICML 2017 and AISTATS 2010, and gave the Breiman Lecture at NIPS 2017. He is interested in the statistical and computational foundations of intelligence, and works on scalable machine learning, probabilistic models, Bayesian non-parametrics and deep learning.

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.  
*KDD 2018, August 19-23, 2018, London, United Kingdom.*

© 2018 Copyright is held by the owner/author(s).

ACM ISBN 978-1-4503-5552-0/18/08.

<https://doi.org/10.1145/3219819.3219941>