Postdoctoral Position to Develop Machine Learning Methods for Analyzing Mathematical Expressions in the Scientific Literature

We have a postdoctoral position to begin in late summer or fall 2015. Supported by a grant from the Alfred P. Sloan Foundation, this is a one year position, with likely renewal to two years.

The position is for researching and developing new methods to analyze mathematical expressions within scientific articles. Our testbed will be the arXiv (http://arxiv.org), a widely used preprint server for hosting and disseminating scientific papers. The goal is to develop advanced mathematical search capabilities using statistical machine learning algorithms to extract representations of the knowledge implicit in the scientific literature. Progress on this research problem has the potential to impact many scientific areas.

We are looking for a postdoctoral scholar with a PhD in Computer Science, Statistics, or a related field. An ideal candidate has strong computational skills and knowledge of statistical machine learning, with fluency in at least a few of the following areas (and willingness to learn about the others):

- graphical models and approximate posterior inference
- collaborative filtering and recommendation systems
- word embeddings
- recurrent neural networks
- statistical translation
- Bayesian nonparametric models
- probabilistic topic models
- matrix factorization
- statistical language modeling

This a joint project between the University of Chicago and Columbia University. The successful candidate will interact with a community of students and faculty at both universities.

To apply, please send a CV to John Lafferty (lafferty@galton.uchicago.edu) and David Blei (david.blei@columbia.edu), and arrange for two letters of reference to be sent to the same addresses.