Large-Scale Machine Learning - Algorithms and Platforms

By: Dr. Rong Yan
April 13, 2010

Abstract:
In recent years researchers are becoming increasingly interested in exploring challenges and new opportunities for developing much larger scale approaches for machine learning and data mining. Many of these computationally-intensive ideas are now becoming practical because of the broader availability of high-speed clusters and the advent of cloud computing. In this talk, I will introduce some of our recent efforts along these directions. This includes our development of more scalable learning algorithms, such as imbalanced boosting, compressed classification, and applications on more advanced distributed computing platforms including Apache Hadoop and a stream processing platform. Finally, I will briefly describe some of the applications of large-scale machine learning in Facebook.

About the Speaker:
Dr. Rong Yan is a currently Research Scientist in Facebook. He was a Research Staff Member in the IBM T. J. Watson Research Center from 2006 to 2009. Dr. Yan received his M.Sc. (2004) and Ph.D. (2006) degree from Carnegie Mellon University's School of Computer Science. His research interests include large-scale machine learning, data mining, advertisement optimization, multimedia information retrieval, and computer vision. Dr. Yan is the leading designer of the automatic video retrieval system that achieves the best performance in the world-wide TRECVID evaluation in 2003 / 2005. He received the Best Paper runner-Up awards in ACM MM 2004 and ACM CIVR 2007. He has received the IBM Research External Recognition Award in 2007. He has served in the NSF proposal review panel and as reviewers for several other research councils. Dr. Yan has authored or co-authored 5 book chapters and more than 60 international conference and journal papers. Dr. Yan has served or is serving as co-chairs for 7 conferences / workshops and as a Program Committee member in more than 35 ACM / IEEE conferences. Dr. Yan gives tutorials and guest lectures at several major conferences and universities. He is an expert reviewer for more than 10 international journals.