Mining with Noise Knowledge: Error Aware Data Mining

Xindong Wu
Department of Computer Science
University of Vermont, Vermont, USA

Abstract

Real-world data are dirty, and therefore, noise handling is a defining characteristic for data mining research and applications. This talk will review existing research efforts on data cleansing and classifier ensembling in dealing with random noise, and then present our recent research on an error aware data mining design to process structured noise. This error aware data mining framework makes use of error information (such as noise level, noise distribution, and data corruption rules) to improve data mining results. Experimental comparisons on real-world datasets will demonstrate the effectiveness of this design.

Biography

Xindong Wu is a Professor and the Chair of the Department of Computer Science at the University of Vermont, USA, and also a Visiting Chair Professor of Data Mining in the Department of Computing at the Hong Kong Polytechnic University, China. He holds a PhD in Artificial Intelligence from the University of Edinburgh, Britain. His research interests include data mining, knowledge-based systems, and Web information exploration. He has published extensively in these areas in various journals and conferences, including IEEE TKDE, TPAMI, ACM TOIS, DMKD, KAIS, IJCAI, AAAI, ICML, KDD, ICDM, and WWW, as well as 14 books and conference proceedings.

Dr. Wu is the Editor-in-Chief of the IEEE Transactions on Knowledge and Data Engineering (by the IEEE Computer Society), the founder and current Steering Committee Chair of the IEEE International Conference on Data Mining (ICDM), the founder and a current Honorary Editor-in-Chief of Knowledge and Information Systems (by Springer), and a Series Editor of the Springer Book Series on Advanced Information and Knowledge Processing (AI&KP). He was Program Committee Chair for ICDM ’03 (the 2003 IEEE International Conference on Data Mining) and Program Committee Co-Chair for KDD-07 (the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining). He is the 2004 ACM SIGKDD Service Award winner, the 2006 IEEE ICDM Outstanding Service Award winner, and a 2005 Chaired Professor in the Cheung Kong (or Yangtze River) Scholars Programme at the Hefei University of Technology sponsored by the Ministry of Education of China and the Li Ka Shing Foundation. He has been an invited/keynote speaker at numerous international conferences including PAKDD-07, IEEE EDOC’06, IEEE ICTAI’04, IEEE/WIC/ACM WI’04/IAT’04, SEKE 2002, and PADD-97.