

From Theory to Practice: A Utility-based Methodology for Mining Object-relational Databases

Massive object-relational databases, which are omnipresent in domains such as computational biology, law enforcement and environmental impact studies, amongst others, bring new challenges to the data mining community. There is an urgent need for novel algorithms and solutions to assist domain experts and data mining novices to understand these vast resources. Such users require direct, transparent access to their large-scale object-relational databases. Furthermore, they require the data mining exercise, and its results, to be evaluated using measures that are of high economic utility, in order to make informed, thought-through decisions. This talk describes our ongoing work, when mining such object-relational databases. We detail our methodology and discuss the results, and lessons learned, when applying these techniques to the CAESARTM anthropometric database and the Protein Data Bank (PDB).

Short Biography

Herna L Viktor is an associate professor at the School of IT and Engineering (SITE), University of Ottawa, Canada. She is a member of the Text Analysis and Machine Learning (TAMALE) group and the leader of the Intelligent Decision Support and Data Analysis Lab (IDeAL) at SITE. Her research focuses on the development of new methodologies for the management and data mining of large-scale object-relational databases and data warehouses. The end results of her research have been applied within the Anthropometry, Health Care and Bioinformatics domains. She holds a Ph. D. in Computer Science from the University of Stellenbosch, which she received in 1999, has published more than 120 international journal and conference articles and is on a number of international programme committees. Her research is sponsored by the Canadian National Science and Engineering Research Council (NSERC), the Canada Foundation for Innovation (CFI) and the Ontario Innovation Trust (OIT).