

# Interestingness in Data Mining

Howard Hamilton<sup>1</sup>

Department of Computer Science, University of Regina,  
3737 Wascana Parkway, Regina, SK, Canada S4S 0A2

**Summary.** Interestingness measures play an important role in data mining regardless of the kind of patterns being mined. These measures are intended for selecting and ranking patterns according to their potential interest to the user. Good measures also allow the time and space cost of the mining process to be reduced. Measuring the interestingness of discovered patterns is an active and important area of data mining research. Although much work has been conducted in this area, so far there is no widespread agreement on a formal definition of interestingness in this context. Based on the diversity of definitions presented to date, interestingness is perhaps best treated as a broad concept, which emphasizes conciseness, coverage, reliability, peculiarity, diversity, novelty, surprisingness, utility, and actionability. This presentation reviews interestingness measures for rules and summaries, classifies them from several perspectives, compares their properties, identifies their roles in the data mining process, gives strategies for selecting appropriate measures for applications, and identifies opportunities for future research in this area.

---

<sup>1</sup>Howard J. Hamilton is a professor in the Department of Computer Science at the University of Regina, Regina, Canada, where he has served since 1991. He received his B.Sc. and M.Sc. in Computational Science from the University of Saskatchewan, and his Ph.D. in Computing Science from Simon Fraser University. He is the Director of University of Regina's Institute for Computational Discovery. His research interests include knowledge discovery in databases, machine learning, applying artificial intelligence to computer animation and computer games, and temporal representation and reasoning. He is a co-author of *Knowledge Discovery and Measures of Interest* and co-editor of four other books, including *Quality Measures for Data Mining*, Springer, 2007. Some issues investigated in knowledge discovery concern interestingness, peculiarity, utility based measures, domain generalization graphs, share-based itemsets, and visualization of results.