Tutorial – 3

Visual Knowledge Discovery and Machine Learning

Prof. Boris Kovalerchuk

Dept. of Computer Science, Central Washington University, USA
http://www.cwu.edu/~borisk

Objectives:

The goal of this tutorial is presenting new visual knowledge discovery and machine learning methods that allow making knowledge discovery and predictive models more effective and rigorous. Specifically we focus on learning tasks of classification and clustering of n-D data using lossless visual representation of n-D data as graphs.

Content and benefits:

Interactive Visual Knowledge Discovery and Machine Learning is a way of enhancing both analytical and visualization methods for discovering hidden patterns in multidimensional data. The fundamental challenge for visual discovery in multidimensional data is that we cannot see n-D data with the naked eye and need visual analytics tools ("n-D glasses"). This challenge starts at 4-D. Often multidimensional data are visualized by non-reversible, lossy dimension reduction methods such as Principal Component Analysis (PCA). While these methods are very useful, they can eliminate important information critical for knowledge discovery in n-D data before starting discovering n-D patterns. Therefore, the expansion of the class of reversible lossless visualization methods is important. The hybrid methods that combine such reversible methods with non-reversible Visualization and Machine Learning methods open new wide opportunities for knowledge discovery in n-D data. These lossless displays are important because of the abilities:

    (1) to restore all attributes of each n-D data point from these graphs,
    (2) to leverage the unique power of human vision to compare in parallel the hundreds of their features, and
    (3) to speed up the selection of an appropriate n-D data classification model.

The presenter will review and compare reversible and non-reversible visual knowledge discovery methods such as General Line Coordinates, PCA, and Multidimensional Scaling, manifolds and others. The successful real-world applications will be presented along with a discussion on how to apply these methods to enhance research in multiple domains. The presenter will use relevant material from his books, including "Visual Knowledge Discovery and Machine Learning" (Springer, 2018) and his recent tutorials at the IEEE and HCI International Conferences and at the e-Science Institute of University of Washington.

The target audience of this tutorial includes graduate students, scientists and practitioners. The benefits from this tutorial for graduate students and researchers are to become familiar with these new developments and new opportunities to enhance their own research inspired by these methods. For practitioners the benefits are in the opportunities to apply these methods to the real world tasks.

Bio Sketch of Presenter:

[Dr. Boris Kovalerchuk](#) is a professor of Computer Science at Central Washington University. USA. His publications include three books "Data Mining in Finance" (Springer, 2000), Visual and Spatial Analysis (Springer, 2005), and "Visual Knowledge Discovery and Machine Learning" (Springer, 2018), a chapter in the Data Mining Handbook and over 170 other publications. His research interests are in visual analytics, data mining, machine learning, uncertainty modeling, data fusion, relationships between probability theory and fuzzy logic, image and signal processing. Dr. Kovalerchuk has been a principal investigator of research projects in these areas supported by the US Government agencies. He served as a senior visiting scientist at the US Air Force Research Laboratory and as a member of expert panels at the international conferences and panels organized by the US Government bodies.

Boris Kovalerchuk, Ph.D.
Professor
Dept. of Computer Science
Central Washington University
Ellensburg, WA 98926-7520, USA
Phone: 509 963-1438
Fax: 509 963-1449
borisk@cwu.edu
http://www.cwu.edu/~borisk