

CORE COMPETENCY WORKSHOP SERIES

# BERD Data Mining in Health Sciences

The CTSI BERD Data Mining in Health Sciences Series is designed to turn attendees into more educated consumers of statistics specifically with regards to data mining. Attendees will gain an overview of data mining methodology as well as details on classification and clustering techniques with applications in health science fields.

This seminar series is designed for health science researchers interested in learning the field of data mining. The talks are designed for an audience with a quantitative mindset but with only a rudimentary knowledge in statistics and probability. Our goal is to educate health science researchers on when data mining techniques are appropriate and to understand the challenges and pitfalls inherent within data mining.

Learners can also earn a certificate of completion upon attending all workshops live, as scheduled, via Zoom.

All sessions will feature faculty from UB's School of Public Health and Health Professions

**DATES:** Wednesdays, April 2022

**TIME:** 4:00 – 6:00 PM

**LOCATION:** Online via Zoom

**REGISTER NOW:** <https://docs.google.com/forms>

For more information, contact [scholar1@buffalo.edu](mailto:scholar1@buffalo.edu) or (716) 829-6268.

## WORKSHOPS

April 13, 2022

**Series Introduction**

**4:00-4:15 PM: Jeffrey Miecznikowski, PhD**  
Associate Professor, Department of Biostatistics

**Overview of Data Mining Methods**

**4:15-6:00 PM: Guan Yu, PhD**  
Assistant Professor, Department of Biostatistics

- Understand common supervised and unsupervised learning problems
- Learn some basic regression techniques

April 20, 2022

**Classification Methods for Data Mining**

**4:00-6:00 PM: Saptarshi Chakraborty, PhD**  
Assistant Professor, Department of Biostatistics

- Understand the basic concept of classification
- Overview some well-known classification methods routinely used in modern-day data mining

April 27, 2022

**Clustering Methods for Data Mining**

**4:00-6:00 PM: Jeffrey Miecznikowski, PhD**  
Associate Professor, Department of Biostatistics

- Understand how to compute similarity of two objects
- Understand the paradigm of most clustering algorithms
- Understand the different approaches for choosing an optimal number of clusters

### EVENT SPONSORS:

**UB Clinical and Translational Science Institute**

- Biostatistics, Epidemiology and Research Design (BERD) Core
- Workforce Development Core
- K Scholar Program

This program is supported by the National Center for Advancing Translational Sciences of the National Institutes of Health under award numbers UL1TR001412 and KL2TR001413 to the University at Buffalo. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.