KENNESAW STATE UNIVERSITY

College of Science and Mathematics

Big Data Lecture Series

CLENDENIN CL2007/Wednesday, July 27th 11:30 – 2:00

Lunch will be served

Applied Machine Learning

Abstract: Building representative machine learning models that generalize well on future data requires careful consideration both of the data at hand and of assumptions about the various available training algorithms. Data are rarely in an ideal form that enables algorithms to train effectively. Some algorithms are designed to account for important considerations such as variable selection and handling of missing values, whereas other algorithms require additional preprocessing of the data or appropriate tweaking of the algorithm tuning parameters. Ultimate evaluation of a model's quality requires appropriate selection and interpretation of an assessment criterion that is meaningful for the given problem. This talk will discuss many of the most common issues faced by machine learning engineers and provides guidance for using machine learning algorithms effectively in practice.

Patrick Hall is a senior staff scientist at SAS Institute and an adjunct professor in the Department of Decision Sciences at George Washington University. He provides guidance for SAS and its customers on deriving substantive value from machine learning technologies and designs new data mining and machine learning approaches, focusing on neural networks and clustering. He holds a patent in automated market segmentation. Patrick is the 11th person worldwide to become a Cloudera certified data scientist. He studied computational chemistry at the University of Illinois before graduating from the Institute for Advanced Analytics at North Carolina State University.