LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-ABS-489119

On The Role of Data Mining Techniques in Uncertainty Quantification

C. Kamath

June 20, 2011

USA/South America Symposium on Stochastic Modeling and Uncertainty Quantification
Rio de Janeiro, Brazil
August 1, 2011 through August 5, 2011

LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-ABS-489119

On The Role of Data Mining Techniques in Uncertainty Quantification

C. Kamath

June 20, 2011

USE/South America Symposium on Stochastic Modeling and Uncertainty Quantification
Rio de Janeiro, Brazil
August 1, 2011 through August 5, 2011

LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

IM release number not issued yet

On The Role of Data Mining Techniques in Uncertainty Quantification

C. Kamath

June 17, 2011

USE/South America Symposium on Stochastic Modeling and Uncertainty Quantification
Rio de Janeiro, Brazil
August 1, 2011 through August 5, 2011

On The Role of Data Mining Techniques in Uncertainty Quantification

Chandrika Kamath
Lawrence Livermore National Laboratory
7000 East Avenue, Livermore, CA 94551

Abstract

Techniques from scientific data mining are increasingly being used to analyze and understand data from scientific observations, simulations, and experiments. These methods provide scientists the opportunity to automate the tedious manual processing of the data, control complex systems, and gain insights into the phenomena being modeled or observed. This process of data-driven scientific inference borrows ideas and solutions from a range of fields including machine learning, image and video processing, statistics, high performance computing, and pattern recognition.

In this talk, I will describe our experiences in using scientific data mining techniques for analysis tasks such as classification of orbits in Poincaré plots, characterization and tracking of blobs in plasma, comparison of simulation and experimental data for code validation, analysis of coherent structures in three-dimensional simulations of the Rayleigh Taylor instability, identification of important sensors associated with edge-harmonic oscillations in a tokamak, inverse design, and analysis of data streams. I will describe the tasks involved in these analyses, such as the extraction of structures from the data, the identification of representative features for these structures, dimension reduction, and building predictive and descriptive models. At first glance, data mining and data driven analysis may appear unrelated to stochastic modeling and uncertainty quantification. But, there are commonalities in the techniques used providing the two communities the opportunity to benefit from the problems in one and the potential solutions in the other.

For additional details:
https://computation.llnl.gov/casc/StarSapphire/
https://computation.llnl.gov/casc/sapphire/