CE 519X: Methods for Data-Driven Computational Engineering Research

Course Description:


No prior experiences with statistical programs, machine learning, high-performance computing and coding are needed.

Course objectives (course designed to provide students with):

This multidisciplinary class is designed to provide graduate students with recent advances in computational and statistical methods that will promote data-driven solutions. Graduate students will be able to tackle problems involving complex, large-scale engineering data. In particular, students will be able to:

1. Effectively shrink and represent complex engineering data and real-world populations.
2. Assess the impact of the data uncertainty on complex engineering responses.
3. Predict engineering responses using cutting-edge statistical learning methods such as the generalized additive model and advanced machine learning methods such as deep learning.
4. Cure incomplete engineering data.

Additional course notes:

- Open-source programs for data-squashing, data-curing, and data-prediction will be provided and utilized for computational exercises and practical projects.
- This class will help educate and motivate students to pursue other classes in computer science or statistics for their specific research needs.
Course Outcomes (students should be able to):

1. Squash complex, large-scale engineering data sets and transform them into tangible and statistically robust representatives using a data squashing program.
2. Assess uncertainty with non-standard distributions of practical engineering data.
3. Make advanced statistical prediction models using the generalized additive model and create a machine learning prediction model using deep learning in the R environment.
4. Cure general incomplete data with complex random missing patterns using an R package FHDI.

Course Topics

- Random sampling versus numerical moment matching technique for:
  - Efficient representation of real-world large-scale engineering data
  - Rigorous uncertainty propagation from raw data to engineering system’s responses
- Statistical learning and prediction versus high-precision computer simulations.
- Generalized additive model (GAM) and its applications to complex engineering problems.
- Introduction to deep learning:
  - Deep neural network for complex data sets.
  - Convolutionary neural networks for two-dimensional image data sets.
- Advanced imputation theories.
- Methods for tackling missing data problems.