

Week 2: Diagnostic Metrics and Cross-Validation

Week 3: Feature Engineering and Behavior Detection

Week 4: Knowledge Inference and Knowledge Structures

Getting Started

Lectures

CTAT Assignment Week 4

Assignment due Aug 26, 2015 at 16:00 UTC



In this assignment, you need to apply a set of metrics to two data sets, [Asgn5-posthoc-data-v2-set1.csv](#) and [Asgn5-posthoc-data-v2-set2.csv](#)

These are not real data sets, but are simulated data, generated for the purpose of this assignment.

You can use any tool you want to complete this assignment. This includes Microsoft Excel, and statistical or data mining package you choose, and other tools available on the internet.

The goal of this assignment is to conduct post-hoc methods to determine which of a large set of correlations can be trusted – and in the process to understand if these post-hoc methods are useful in educational data mining.

---