ONE PAGE COURSE OVERVIEW

Course Title: D209 – Data Mining I

Course Description: Data Mining I expands predictive modeling into nonlinear dimensions, enhancing the capabilities and effectiveness of the data analytics lifecycle. In this course, learners implement supervised models—specifically classification and prediction data mining models—to unearth relationships among variables that are not apparent with more surface-level techniques. The course provides frameworks for assessing models' sensitivity and specificity.

Competencies:

a. Classification Data Mining Models: Apply observations to appropriate classes and categories using classification models
b. Predictive Data Mining Models: Implement prediction data mining models to find hard-to-spot relationships among variables.
c. Data Mining Model Performance: Evaluate data mining model performance for precision, accuracy, and model comparison.

Performance Assessment (PA): Task covering the above competencies and involving:

Task 1:
Create a data mining report using one of the following classification methods (*k*-nearest neighbor (KNN) or Naive Bayes).

Task 2:
Create a data mining report using one of the following classification methods (decision trees, random forests or advanced regression (i.e., lasso or ridge regression)

Tools and Techniques: Python or R. (Students can choose).

Resources:

a. Datacamp