Executive Certificate in Data Science
Big Data and Machine Learning for Process Improvement Professionals

Overview

This 7-week program will introduce participants to the basic methods of machine learning and big data, as applied to process improvement challenges. The curriculum combines online coursework, live virtual sessions, and practice with discrete event simulation software.

Participants will learn new analytical techniques to generate powerful insights and then turn those insights into improved process performance.

Who Should Register?

If you are targeting the growing field of data analytics or looking to strengthen your data-driven, decision-making skills, this program will help you achieve your professional goals.

Typical participants include Process Improvement Leaders, Lean Six Sigma Black Belts, Master Black Belts, Process Engineers, and Quality Improvement Professionals. The program curriculum assumes participants have a solid understanding of probability and statistics, linear regression, and process management.

Sessions Start September 2021

This program will meet each Tuesday and Thursday, beginning September 21st and will continue until November 4th. Live sessions will be held via Zoom from 3:00 pm to 4:30 pm (Eastern).

Program fees are $5000 and, in addition to the virtual sessions, include:

- Advanced Process Analytics online course (92 hrs.) – Covers all prerequisites, to refresh your knowledge.
- Process Modeling course (10 hrs.)
- One-year subscription to EngineRoom software with Process Playground

Curriculum

WEEK 1	• **The Data Science Process (1.5 hrs)** – In this opening session, participants will be exposed to the basic process of working with data. This session will discuss the challenges and ideas related to those steps and lay a foundation for the remainder of the workshop. • **Introduction to R 1 (1.5 hrs)** – The R programming language is a cornerstone of modern data science. In this hands-on session, participants will explore the basics of R and the Tidyverse.
WEEK 2	• **Introduction to R 2 (1.5 hrs)** – Students will continue to work with the R language and powerful data visualization capabilities of the ggplot2. • **Planning for Data Quality (1.5 hrs)** – Data quality is a critical consideration. Practical tips to avoid data quality problems through better planning for data collection and structure will be covered.
WEEK 3	• **Simulation and Bootstrapping (1.5 hrs)** – In this session, participants will examine tools that allow them to model and study random processes. These tools will be used to help participants understand machine learning in the data science process. • **Introduction to Machine Learning Models (1.5 hours)** – Machine learning models are powerful tools for eliciting the story behind the data. This session will give an overview of supervised and unsupervised machine learning and demonstrate methods of evaluating these algorithms.
WEEK 4	• **Machine Learning for Classification (1.5 hrs)** – This session will help continue the exploration of supervised machine learning models. This session will introduce programming machine learning algorithms for classification. Students will explore the fundamentals of decision trees and k-nearest neighbors algorithms and discuss evaluating these algorithms. • **Ensemble Methods in Machine Learning (1.5 hrs)** – Ensemble methods combine multiple machine learning models to refine and improve model performance. This session will expose students to these ensembles and techniques, such as random forests.
WEEK 5	• **AI and Neural Networks 1 (1.5 hrs)** – Neural networks have broad applications including much of what we think of as Artificial Intelligence (AI). In this session, participants will see the basics of training and using neural networks. • **AI and Neural Networks 2 (1.5 hrs)** – This session will continue exploration of neural networks and their application to complex tasks like image recognition.
WEEK 6	• **Simulation with Process Playground (1.5 hrs)** – This session will introduce participants to the powerful techniques for process modeling using discrete event simulation software. • **Unsupervised Learning (1.5 hrs)** – Machine learning is adept at finding hidden groups and relationships in data. This session allows participants to explore algorithms for unsupervised learning such as hierarchical clustering and k-means clustering.
WEEK 7	• **Using Process Playground (1.5 hrs)** – This session will give participants more hands-on experience putting Process Playground to use in process simulation. • **Applying Machine Learning in Process Improvement (1.5 hrs)** – This final session will allow students to look back at the machine learning tools and apply them to a process improvement scenario.