CERTIFICATE IN **DATA SCIENCE**

OVERVIEW

Data Science is impacting many areas of science, engineering, and industry; from analyzing troves of weather data to modeling traffic patterns to processing millions of online customers, it is the enormous data which is creating new opportunities and challenges.

To tackle these challenges, one must have the training to store, manage, process and analyze data at these scales. But the challenges are beyond scale alone, the complexity of the data requires new powerful analytical techniques. Finally, it is crucial to have skills in communicating and interpreting the results of this analysis. A person trained in all of these skills is a **big data scientist.**

CORE CLASSES Must take at least three of the following.	
CS 6140	Data Mining
CS 6350	Machine Learning
CS 6530	Database Systems /or/ 5530 Database Systems
CS 6630	Visualization

GET UP TO SPEED   Optionally take one.

COMP 5005	Programming for Engineers
COMP 5360	Introduction to Data Science
COMP 5960	Foundations of Data Analysis

ELECTIVE CLASSES   These classes fulfill the additional class requirements.

ALGORITHMICS

CS 6150	Graduate Algorithms
CS 6160	Computational Geometry
CS 6170	Computational Topology
CS 6180	Clustering

ANALYTICS

CS 6190	Probabilistic Learning
CS 6210	Advanced Scientific Computing
CS 6300	Artificial Intelligence
CS 6340	Natural Language Processing
CS 6630	Vis for Data Science
CS 6550	Information Retrieval
CS 6961	Structured Prediction

MANAGEMENT

CS 6230	High-Performance Computing and Parallelization
CS 6235	Parallel Programming for GPUs/Many Course/Multi-Cores
CS 6480	Advanced Computer Networks
CS 6490	Network Security
CS 6530	Advanced Database Systems
CS 6963	Distributed Systems

Additional substitutions may be approved by the director on a case-by-case basis.

   Students must complete five classes (15 credit hours) with a B or better. Other than the core classes, and the optional "Get up to speed" classes, the additional classes can include any other graduate level classes approved at the discretion of the director. The elective classes are pre-approved to fulfill the requirements, but many other (often more sporadically offered) classes are available. There can be up to one class from a Data Domain (at least 3 credit hours), which are typically a course in another department uses data science techniques within the context of a specific domain.