CERTIFICATE IN **BIG DATA**

OVERVIEW

Big Data 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 classes from at least 4 of the 5 lines.	
CS 6140	Data Mining
CS 6150	Advanced Algorithms
CS 6350	Machine Learning
CS 6530	Database Systems /or/ 5530 Database Systems
CS 6630	Visualization

ELECTIVE CLASSES These classes fulfill the 5th class requirement.	

ALGORITHMICS

CS 6160	Computational Geometry
CS 6170	Computational Topology
CS 7960	Models of Computation for Massive Data

ANALYTICS

CS 6190	Probabilistic Modeling
CS 6210	Advanced Scientific Computing
CS 6300	Artificial Intelligence
CS 6340	Natural Language Processing
CS 6640	Image Processing

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

Additional substitutions may be approved by the Data Management and Analysis Track Director on a case-by-case basis.

Students must complete 5 classes (15 credit hours) with a B or better. At least 4 classes must be among the CORE classes (it is suggested to take all 5). The 5th classes can be any other graduate level classes approved by the Data Management and Analysis Track director. The ELECTIVE classes are pre-approved to fullfill this requirement, but many other (often more sporadically offered) classes are available.