Large high-dimensional data sets are used in an increasing number of scientific fields and are characterized by considerable amount of information content that needs to be processed. Application areas include bioinformatics, remote sensing (e.g. hyperspectral datasets), homeland defense (e.g. face recognition, epidemiology) large-scale physics simulations, and dynamics on complex networks (e.g. internet traffic analysis, urban population dynamics).

The course will provide an overview of the basic processing techniques available for large data sets. In addition to extracting relevant information, such data also need to be transmitted, stored, and compressed. Finally, the course will emphasize computational environments that yield to speed-ups in processing such as parallel, grid and distributed environments.

The course sequence includes:
- intro to general pattern recognition
- study of the HQD types
- collection and processing HQD
- feature extraction in large data sets
- parallel processing of large data sets
- distributed and grid computing and large data sets

The course will use various outside references provided by the instructor, with no textbook required. The class includes a term project on PD topic of choice. The course will be offered using Tablet PCs in classroom.

The course meets Thursday 5:30-8:00pm In Richardson Hall 117.

For more information, contact:

Stefan Robila
robilas@mail.montclair.edu
973-655-4230

<<<<<<<<<This course is open to both undergraduate and graduate students>>>>>>>