One of the most important scientific events of the 21st century is the Human Genome Project, which brings the complete sequence of human being into hands of anybody who has a computer. In this course, we will introduce principles and algorithms from statistics, machine learning, and pattern recognition to address exciting biological problems such as gene discovery, gene function prediction, gene expression regulation, diagnosis of cancers, etc. This course will take a case-study approach to current topics in bioinformatics. A series of projects emphasize real-life data, hands-on analysis, and collaboration. Course projects involve C/C++ and/or PERL programming.

Prerequisite: basic skills of C/C++/Java programming, PERL programming skill is a plus but not required.  
Keywords: Genes, Genome sequences, gene expression regulation, machine learning, pattern recognition, disease diagnosis

Textbooks (Tentative): Introduction to Computational Genomics: A Case Studies Approach. (Check it at Amazon)

Tentative Topics:

- Introduction to DNA, RNA, proteins and central dogma of molecular biology
- Sequence alignment algorithms and sequence database retrieval
- Gene finding: how to find a disease gene?
- Gene/protein function prediction: who does what?
- Hidden Markov Models for sequence pattern recognition: math talks
- DNA regulatory motif analysis: How the billions of cells are programmed to work?
- Phylogenetic analysis: trace the history of human being
- Whole genome comparison: How different between you and chimpanzee?
- Introduction to Structural bioinformatics: protein folding
- Microarray based Gene expression analysis: the real story of cells
- Case studies: find the bad genes, predict the diseases

Job market for bioinformatician is expanding dramatically in biotechnology and pharmaceutical industries. It is fun to understand life with your laptop computer!