Course meeting times and location:

- 227 East 30th Street (Translational Research Building), TBD
- Friday afternoon (TBD)

Faculty contact info & office location:

Alexander Statnikov, Ph.D.
- 227 East 30th Street (Translational Research Building), 7th floor, Office #736
- Email: Alexander.Statnikov@med.nyu.edu
- Office phone: 212-263-3641
- Cell phone: 615-545-3685

Constantin F. Aliferis, M.D., Ph.D., F.A.C.M.I.
- 227 East 30th Street (Translational Research Building), 7th floor, Office #743
- Email: Constantin.Aliferis@nyumc.org
- Office phone: 212-263-5281

Prerequisites: Foundational coursework in Mathematics and Computer Science is required. Specifically, students are required to take the following courses prior to registering: Calculus (2 semesters), Linear Algebra, Algorithms and Data Structures, and Introduction to Programming. In some circumstances, students who do not meet the above requirements can obtain permission of the instructor to take the course.

Course topics by weeks:

1. Overview of classes of machine learning problems
2. Basic principles of supervised learning (classification) and necessary mathematical concepts
3. Support vector machines for binary classification
4. Other methods for binary classification
5. Basic principles of statistical machine learning
6. Model selection and accuracy estimation
7. Multicategory classification
8. Regression
9. Novelty detection
10. Clustering
11. Variable selection (week 1)
12. Variable selection (week 2)
13. Machine learning for genomic data
14. Machine learning for text data
15. Machine learning for clinical data

**Course materials:**

**Required books:**


**Optional book:**


**Computer tools and software:**

- Matlab (Matlab installation instructions are attached). Required toolboxes: Bioinformatics, Statistics, Neural Network, Optimization.
- LibSVM library for Matlab (http://www.csie.ntu.edu.tw/~cjlin/libsvm/)
- LibLinear library for Matlab (http://www.csie.ntu.edu.tw/~cjlin/liblinear/)

**Learning objectives:** Understand methods and techniques for machine learning.

**Assignments and assessment:**

- Readings assignments (textbooks, articles)
- Practical take-home assignment #1: Intro to Matlab
- Practical take-home assignment #2: Building classification models and estimating their accuracy
- Practical take-home assignment #3: Comparing various algorithms for classification
- Weekly homework assignments geared towards understanding of algorithms
- Midterm examination
- Final examination

Each practical take-home assignment contributes 10% of the grade (30% for practical assignments in total), each examination contributes 30% of the grade (60% for exams in total), homework assignments contribute a total of 10%.

**Course Assessment:** This course can be taken for a letter grade or audited.