# [BMED-4813R/8813-BHI] Syllabus

**[Biomedical and Health Informatics]**

## Instructor Information

**Instructor**
Dr. May D. Wang

**Email**
maywang@gatech.edu

## General Information

### Description

This course intends to introduce Biomedical Big Data and Biomedical and Health Informatics (BHI) to senior undergrad and graduate students of BME. It covers: (1) informatics needs driven by current biomedicine and health care (e.g., cancer, cardiovascular disease, population and public health such as COVID-19 Pandemic etc.); (2) informatics challenges and common methodologies; and (3) progress and opportunities in BHI. In 2004, NIH defined 'Biomedical Informatics' as a broad area of research that includes bioinformatics (computational molecular biology), biomedical imaging informatics, clinical informatics (e.g., personal health records [PHR], and electronic health records [EHR]), and public health informatics (i.e., population public health challenges such as infectious diseases -- COVID-19, Influenza, AIDs outbreaks in epidemiology). As summarized by the US National Academy of Engineering, 'Health Informatics' is one of 14 Grand Challenges for the 21st Century.

### Prerequisites

This course is suitable for senior undergraduate students and entry-level graduate students. We will provide review of basic statistics and data processing concepts (including 1-D, and multi-dimensional signals), and recommend that students know at least one programming language (e.g., MATLAB, C, R, Python, Perl, Java, and PHP etc.) to assist the students in getting the most out this course through homework practice, and projects. Strong self-motivation to gain knowledge, skills, and problem solving ability in this fast growing field of Biomedical Big Data is necessary!

### Course Goals and Learning Outcomes

This course is a course that integrates traditional classroom lecturing and problem-based learning. Students will learn Biomedical and Health Informatics techniques by solving a real-world problem (real-world biomedical or health data are provided) beyond classroom lectures. They will work in teams to learn skills and to apply critical thinking and knowledge inquiry towards solving real-world biomedical and health problems.

Health includes personal and population. For population health, in lieu of recent COVID-19 pandemic outbreak, students will use population big data for infectious disease spread modeling, and will use personal big data in sensing, personal health records, and clinical records for disease screening/diagnosis, treatment, and prognosis decision making. Besides teaching systems models to model population health, this course will also go through the typical flow of clinical information with the following steps: (1) data acquisition, storage, and curation; (2) data analysis, mining, and visualization; (3) modeling; (4) decision support; and (5) delivery (point-of-care access). Example topics are organized as below:

- Review of Basic Biostatistics Methods and Tools
- Introduction of Basic Pattern Recognition and Decision Making Concepts
- Introduction of Commonly Used Data Mining Methodologies
- Introduction of Performance Evaluation Metrics
- Tutorial on Python, MATLAB, and Deep Learning Library
- Problem Solving Pipeline for Clinical Decision Support
- Review of Basic Digital Signal Processing Techniques and Tools
- Review of Systems Modeling (Agent-based, Compartmental-based etc.)

- Application 1: Systems Modeling for Infectious Disease Spreading such as COVID-19.
- Application 2: Electrical Medical Record Data Quality Control and Health Outcome Prediction
- Application 3: High-Throughput –Omic Next-Generation Sequencing (NGS) Bioinformatics (Cancer or Neural Disease Diagnosis)

- Application 4: Physiological Informatics for Health Monitoring (e.g. COVID-19 Patient Home-Based and Hospital-Centered Monitoring,

## Course Requirements & Grading

| **Team-Based (50%)** | |
|---|---|
| Study and Present Peer-Reviewed Paper | 10% |
| Write a Literature Critique Article | 30% |
| Rapid Fire Style Presentation of Proposal/Critique | 10% |
| **Individual-Based Assessment (40%)** | |
| Homework (including Programming) | 40% |
| **Team-Member Effort Peer Evaluation** | **5%** |
| **Peer Evaluation of Rapid Fire** | **5%** |

### Grading Policies

a) Late assignments will not be graded and will receive a score of 0.
b) All grades become final one week after they are returned. Any appeals to a returned grade must be made in writing by email to the professor within this one-week window to be considered.
c) The lowest homework grade will be dropped.

## Course Expectations & Guidelines

### Course Website and Other Classroom Management Tools

Canvas will be used for announcements, assignments, and discussions.

### Course Email Format

(1) All private email communications with Professor Wang should have Subject Line Format:

**CourseSession (Summer2021-BMED4813): (your first name, your last name) - (subject)**

(2) All general technical questions, please use Canvas for others to view, OR to send to Professor and copy TA to get timely response.

### Academic Integrity

Georgia Tech aims to cultivate a community based on trust, academic integrity, and honor. Students are expected to act according to the highest ethical standards. For information on Georgia Tech's Academic Honor Code, please visit http://www.catalog.gatech.edu/policies/honor-code/ or http://www.catalog.gatech.edu/rules/18/.

Any student suspected of cheating or plagiarizing on a quiz, exam, or assignment will be reported to the Office of Student Integrity, who will investigate the incident and identify the appropriate penalty for violations. Furthermore, all students are expected to report, to the instructor, all incidents or suspected incidents of academic misconduct in this class.

### Accommodations for Students with Disabilities

If you are a student with learning needs that require special accommodation, contact the Office of Disability Services at (404)894-2563 or http://disabilityservices.gatech.edu/, as soon as possible, to make an appointment to discuss your special needs and to obtain an accommodations letter. Please also e-mail me as soon as possible in order to set up a time to discuss your learning needs.

### Attendance

Students are responsible for all material covered in class, including but not limited to, lecture topics, handouts, assignments, and announcements.

### Collaboration & Group Work

Students **CANNOT** collaborate on individual-performance-required assignments such as exams. Collaborative discussion is permitted for homework with final submission as an individual submission. For the team-based projects, students will work together in teams of 2(+1) members depending on course size; the contribution of individual team members will be evaluated by peers, and count as five (5) points towards the final grade.

### Student-Faculty Expectations Agreement

At Georgia Tech we believe that it is important to strive for an atmosphere of mutual respect, acknowledgement, and responsibility between faculty members and the student body. See http://www.catalog.gatech.edu/rules/22/ for an articulation of some basic expectation that you can have of me and that I have of you. In the end, simple respect for knowledge, hard work, and cordial interactions will help build the environment we seek. Therefore, I encourage you to remain committed to the ideals of Georgia Tech while in this class.

## Course Materials

### Course Texts

No single text will cover the topics. Copies of papers and chapters from different books or journals will be provided as reading references.

- *Edward H. Shortliffe and James J Cimino (editors),* Biomedical Informatics: Computer Applications in Health Care and Biomedicine (Health Informatics), 4th Edition, Oct. 2013.
- *Mark L Braunstein,* Health Informatics in the Cloud, 2013.
- *James H McClellan, Ronald W Schafer, and Mark A Yodar*, Signal Processing First, 2003.
- *Brani Vidakovic,* Statistics for Bioengineering Sciences: With MATLAB and WinBUGS Support, 2011.
- *Richard O Duda,* Pattern Classification with Computer Manual, 2nd Edition, 2014.
- *Edward R Tufte,* The Visual Display of Quantitative Informatics, 2001.
- *Christopher Bishop,* Pattern Recognition and Machine Learning, 2007.

## Topics Covered

(A) Motivation and Basic Review in Big Biomedical Data Era
- What is Biomedical Big Data (5V, BD2K, Privacy, Security)?
- What is Biomedical and Health Informatics (AMIA and NIH definition)?
- Review of Basic Biostatistics Methods and Tools
- Introduction of Basic Pattern Recognition and Decision Making Concepts (Learning, Distance, Cross Validation, External Validation, Predictive Model Selection)
- Introduction of Commonly Used Data Mining Methodologies (Clustering vs Classification, Logistic Regression, K-Means, Bayesian Classifier, LDA, KNN, SVM, NN etc)
- Introduction of Performance Evaluation Metrics (e.g. Accuracy, Precision, MCC, Sensitivity, Specificity, AUC, FDR etc.)
- Problem Solving Pipeline for Clinical Decision Support
- Clinical Decision Support

(B) Predictive Modeling Application: Public Health Informatics and Systems Medicine
- Overview and Systems Modeling (MATLAB SIMULINK)
  - ODE-Based Model
  - Agent-Based Model
- Parameter Estimation for SEIR, IDEA, etc.
- Population Modeling Case Studies

  - Google Flu
  - COVID-19 Outbreak Prediction
- Simulation of Biological Systems: Biochemical and Cell/Organ Models

(C) Bioinformatics Application: High-Throughput –Omic and Next-Generation Sequencing (NGS)

- Data Acquisition Technologies: Microarrays and Sequencing
- Microarray Quality Control and Quantification Methods
- NGS: RNA-Seq Mapping, Quantification, and Normalization Methods
- Molecular Signature (Biomarker) Identification Case Studies:
  - Cancer (The Cancer Genome Atlas)
  - Asthma
  - Genomics Characterization of Infectious Disease (e.g. Ebola)

(D) Health/Clinical Informatics

- Personal Health Record (PHR) and Electronic Health Record (EHR) Data Analytics
  - Quality Control
  - Heterogeneous Data Mining
- Case Studies
  - Electrical Medical Record Data Quality Control and Health Outcome Prediction
- Emerging Topics

(E) Medical Imaging Informatics for Disease Diagnosis

- Translational Informatics Pipeline
  - Quality Control (e.g. Data Normalization)
  - Information Extraction (Pixel-Based or Object-Based Image Feature Extractsion)
  - Knowledge Modeling
  - Clinical Decision Making
  - Feedback and Validation