INFO-Bioinformatics

INFO 501. Biomed Informatics Research. 3 Hours.
Biomedical Informatics Research. Biomedical informatics is the art and science of collection, representation and analysis of information for the purpose of improving human health. Informatics applications span the spectrum from molecular (bioinformatics) to organism (clinical informatics). This course will examine the scientific field that underlies the development of tools and methods applied to the biomedical domain. The course will include lectures, readings from a textbook and journal papers, a term paper reviewing some area of informatics research, and a final examination. It is intended for students who are studying applied areas of informatics (including Health Informatics and Nursing Informatics) as well as students who would like to explore the possibility of an informatics research career.

INFO 601. Introduction to Bioinformatics. 3 Hours.
Introduction to bioinformatics and computational biology, with emphasis on concepts and application of informatics tools to molecular biology. It covers biological sequence analysis, gene prediction, genome annotation, gene expression analysis, protein structure prediction, evolutionary biology and comparative genomics, bioinformatics databases, cloud computing, basic R-based data analysis, simple programming skills using Perl, Linux/Unix environment and command lines, visual analytics, and social/legal aspects of open science. It will have a class research project component.

INFO 602. Algorithms in Bioinformatics. 3 Hours.
This course introduces various fundamental algorithms and computational concepts for solving questions in bioinformatics and functional genomics. These include graph algorithms, dynamic programming, combinatorial algorithms, randomized algorithms, pattern matching, classification and clustering algorithms, hidden Markov models and more. Each concept will be introduced in the context of a concrete biological or genomic application. A broad range of topics will be covered, ranging from genome annotation, genome reconstruction, microarray data analysis, phylogeny reconstruction, sequence alignments, to variant detection.
Prerequisites: INFO 601 [Min Grade: C]

INFO 603. Biological Data Management. 3 Hours.
The introduction of biological data management concepts, theories, and applications. Basic concepts such as relational data representation, relational database modeling, and relational database queries will be introduced in the context of SQL and relational algebra. Advanced concepts including ontology representation and database development workflow will be introduced. Emerging big data concepts and tools, including Hadoop and NoSQL, will be introduced in the context of managing semi-structured and unstructured data. Application of biological data management in biology will be covered using case studies of high-impact widely used biological databases. A class project will be required of all participants.
Prerequisites: INFO 601 [Min Grade: C]

INFO 604. Next-generation Sequencing Data Analysis. 3 Hours.
This course is aimed to equip participants with the essential knowledge and skills required to begin analyzing next-generation sequencing data and carry out some of the most common types of analysis. The topics covered in-depth during this course are the analysis of RNA-Seq, ChIP-Seq data, ATACseq data, and Single-cell data, with an optional Variant Calling session. The sessions will also include Introduction to next-generation sequencing (NGS) technologies, common NGS data analysis issues, applications of sequencing technologies, introduction to bioinformatics file formats (e.g. FASTQ, bam, bed) and bioinformatics toolkits. At the end of this course, participants will have the expertise to perform these data analysis independently.
Prerequisites: INFO 601 [Min Grade: C]

INFO 610. Programming with Biological Data. 3 Hours.
This course is a 4-week session, class meets 4 days a week from 9:00am-11:30am. Class will meet at Tinsley Harrison Tower, room 144. This course provides students necessary bioinformatics programming and data skills using Linux, MySQL and Python. Emphasis will be placed on best practices for conduct of reproducible research. The focus will be on practical computing, programming and data management concepts that can be applied to bioinformatics and data analysis problems.

INFO 611. Intermediate Statistical Analysis I. 3 Hours.
Students will gain a thorough understanding of basic analysis methods, elementary concepts, statistical models and applications of probability, commonly used sampling distributions, parametric and non-parametric one and two sample tests, confidence intervals, applications of analysis of two-way contingency table data, simple linear regression, and simple analysis of variance. Students are taught to conduct the relevant analysis using current software such as the Statistical Analysis System (SAS).

INFO 612. Visual Analytics for Bioinformatics. 3 Hours.
In this course, we will explore the use of visualization techniques as a concise and effective way to help analyze, understand, interpret and communicate complex biological data. Principles of design, visual rhetoric/communication, and appropriate usage will be introduced. We will cover representation of different data types, concentrating on those generated by data-rich platforms such as next-generation sequencing applications, flow/mass cytometry, and proteomics, and will discuss the use of visualization techniques applied to assessing data quality and troubleshooting. Various topics including dimension reduction, hierarchical visualizations, unsupervised learning, graph theory, networks/layouts and interactivity will be discussed. We will review the algorithmic underpinnings of various methods that lead to their appropriate and effective use. Finally, we will review a variety of genomics/bioinformatics-related visualization tools that are available. We will use Matlab throughout the course to create beautiful and effective visualizations.
Prerequisites: INFO 603 [Min Grade: C]

INFO 651. Systems Biomedicine of Human Microbiota. 3 Hours.
The human microbiota is the collection of microorganisms (bacteria, archaea, fungi and viruses) that reside within human tissues and biofluids. Such resident microorganisms compose the majority of cells in human bodies and are key contributors to human development, health, and disease. However, most studies focus on genomics and microbiome statistical representations alone, while spatial-temporal analysis, multi-source data integration and modeling are necessary to predict and understand interactions between microorganisms, human hosts, and the environment. This course will highlight state-of-the-art microbiome/microbiota research and provide essential training in mathematical, computational and systems biology to derive integrative and predictive models of microbiota-host interactions in the context of human health and disease.
Prerequisites: INFO 601 [Min Grade: C] and (MA 560 [Min Grade: C] or BME 670 [Min Grade: C])

INFO 662. Biomedical Applications of Natural Language Processing. 3 Hours.
Students will be introduced to Natural Language Processing (NLP) including core linguistic tasks such as tokenization, lemmatization/stemming, POS tagging, parsing and chunking. Applications will focus on Deep Learning methods using pytorch with a focus in information extraction including Named Entity Recognition, semantic role labeling, word sense disambiguation, normalization, summarization, question answering and text classification. Applications and data will have a biomedical focus, but no biology or medical background is required.

INFO 671. Clinical Informatics Seminar I. 1 Hour.
For master's student only. Students will learn how to prepare, present, and critique research presentations in clinical informatics by attending seminar presentations made by presenters. Seminars are presented by graduate students, faculty, visitors, or online speakers. Students must show evidence of prior preparation, active participation, and documented comprehension of the topics.
Prerequisites: INFO 501 [Min Grade: C]

INFO 672. Clinical Informatics Seminar II. 1 Hour.
For master's student only. Students will learn how to prepare, present, and critique research presentations in clinical informatics by attending seminar presentations made by presenters. Seminars are presented by graduate students, faculty, visitors, or online speakers. Students must show evidence of prior preparation, active participation, and documented comprehension of the topics.
Prerequisites: INFO 671 [Min Grade: C]

INFO 673. Clinical Informatics Journal Club. 0-1 Hours.
Students will learn how to read, present, and critique primary research publications in clinical informatics. Journal club participants will present high-impact recent journal publications selected by course instructors and learn how to read the paper, write critiques, and organize analysis insights into review papers. Students must show evidence of prior preparation prior to journal clubs and write critiques to show comprehension of the topics throughout the semester.

INFO 680. Implementation and Evaluation of Clinical Systems. 3 Hours.
Health information technology (HIT) tools such as Electronic Health Records (EHRs) are used to facilitate management of patient care data, to computerize clinical workflows, and to support health professionals in their medical decision making process. As a result of the U.S. Federal Government incentive program known as Meaningful Use, EHRs have been adopted on a national scale and are now used in almost every health care organization across the country. Although the literature exploring the impact of HIT adoption and use has also increased, previous studies have produced mixed results, leaving unanswered questions as to the impact of HIT on quality of care, patient safety, and health care providers' productivity. In this course, students will be introduced to project management tools and techniques commonly used for managing implementation of HIT systems as well as research approaches to conduct systematic evaluations of the impact of these systems on health care outcomes and organizations. This foundational course is intended for informatics majors and students in allied fields (e.g., health, biological, or computer sciences) who are interested in exploring implementation methods applicable to HIT systems such as EHRs and their components, as well as quantitative, qualitative, and mixed-methods approaches to conduct evaluations of HIT adoption and use. It is primarily intended for students who will pursue research careers in biomedical informatics and is the third course in a three-part series.
Prerequisites: INFO 697 [Min Grade: C]

INFO 690. Data Mining & Statistical Learning. 3 Hours.
Students will learn to discover and implement meaningful insights and knowledge from data. This course covers major concepts and algorithms of data mining. The course will be taught using the SAS Enterprise Miner program. The final project will demonstrate all the data mining techniques covered in the course and furthermore expose students working with real data. At the end of the course students will be proficient in utilizing data mining techniques to exploit data patterns and behavior, gain insider understanding of the data, and produce new knowledge that healthcare decision-makers can act upon. Furthermore, SAS Certified Predictive Modeler certification exam will be offered at the end of the course. Instructor permission is required.

INFO 691. Bioinformatics Seminar I. 1 Hour.
For master's student only. Students will learn how to prepare, present, and critique research presentations in bioinformatics by attending seminar presentations made by presenters. Seminars are presented by graduate students, faculty, visitors, or online speakers. Students must show evidence of prior preparation, active participation, and documented comprehension of the topics.
Prerequisites: INFO 601 [Min Grade: C]

INFO 692. Bioinformatics Seminar II. 1 Hour.
For master's student only. Students will learn how to prepare, present, and critique research presentations in bioinformatics by attending seminar presentations made by presenters. Seminars are presented by graduate students, faculty, visitors, or online speakers. Students must show evidence of prior preparation, active participation, and documented comprehension of the topics.
Prerequisites: INFO 691 [Min Grade: C]

INFO 693. Bioinformatics Journal Club. 2 Hours.
Students will learn how to read, present, and critique primary research publications in bioinformatics. Journal club participants will present high-impact recent journal publications selected by course instructors and learn how to read the paper, write critiques, and organize analysis insights into review papers. Students must show evidence of prior preparation prior to journal clubs and write critiques to show comprehension of the topics throughout the semester.

INFO 695. Special Topics in Bioinformatics. 3 Hours.
Topics of current research interest, such as metagenomics, microbiome, computational medicine, complex systems, deep learning in biology, artificial intelligence in biomedical, and translational bioinformatics applications. May be repeated as different sections taught by different instructors for credit. Permission of instructor is required.

INFO 696. Biomedical Informatics Methods I. 3 Hours.
Biomedical informatics is the art and science of collecting, representing and analyzing patient and biomedical information and translating insights from the information into better health and new medical discoveries. The spectrum of informatics applications ranges from molecules (bioinformatics) to individuals and populations (clinical and public health informatics). We will examine the scientific field and research methods that form the foundation for biomedical informatics research. The course will include didactics, readings, hands-on tool explorations, and a summative work product. This foundational course is intended for informatics majors and students in allied fields (e.g., health, biological, or computer sciences) who are interested in exploring the field of informatics.

INFO 697. Biomedical Informatics Methods II. 3 Hours.
Biomedical informatics is the art and science of collecting, representing and analyzing patient and biomedical information and translating insights from the information into better health and new medical discoveries. The spectrum of informatics applications ranges from molecules (bioinformatics) to individuals and populations (clinical and public health informatics). We will examine the scientific field and research methods that form the foundation for biomedical informatics research. The course will include didactics, readings, and applications in applying research methods, culminating in a research plan in grant proposal format and review by a mock panel. This foundational course is intended for informatics majors and students in allied fields (e.g., health, biological, or computer sciences) who are interested in exploring the field of informatics. It is primarily intended for students who will pursue research careers in biomedical informatics and is the second course in a two-part series.
Prerequisites: INFO 695 [Min Grade: B] or INFO 696 [Min Grade: B]

INFO 698. Bioinformatics Master's Projects. 1-6 Hour.
Admission to bioinformatics master's program (Plan B: "Project Option") is required. Independent study to conduct bioinformatics research projects, guided by the instructor as the mentor. Permission of instructor and graduate program director is required.

INFO 699. Bioinformatics Master's Thesis Research. 1-6 Hour.
Admission to bioinformatics master's program (Plan A: "Thesis Option") is required.

INFO 701. Introduction to Bioinformatics. 3 Hours.
Introduction to bioinformatics and computational biology, with emphasis on concepts and application of informatics tools to molecular biology. It covers biological sequence analysis, gene prediction, genome annotation, gene expression analysis, protein structure prediction, evolutionary biology and comparative genomics, bioinformatics databases, cloud computing, basic R-based data analysis, simple programming skills using Perl, Linux/Unix environment and command lines, visual analytics, and social/legal aspects of open science. It will have a class research project component.

INFO 702. Algorithms in Bioinformatics. 3 Hours.
This course introduces various fundamental algorithms and computational concepts for solving questions in bioinformatics and functional genomics. These include graph algorithms, dynamic programming, combinatorial algorithms, randomized algorithms, pattern matching, classification and clustering algorithms, hidden Markov models and more. Each concept will be introduced in the context of a concrete biological or genomic application. A broad range of topics will be covered, ranging from genome annotation, genome reconstruction, microarray data analysis, phylogeny reconstruction, sequence alignments, to variant detection.
Prerequisites: INFO 701 [Min Grade: C]

INFO 703. Biological Data Management. 3 Hours.
The introduction of biological data management concepts, theories, and applications. Basic concepts such as relational data representation, relational database modeling, and relational database queries will be introduced in the context of SQL and relational algebra. Advanced concepts including ontology representation and database development workflow will be introduced. Emerging big data concepts and tools, including Hadoop and NoSQL, will be introduced in the context of managing semi-structured and unstructured data. Application of biological data management in biology will be covered using case studies of high-impact widely used biological databases. A class project will be required of all participants.
Prerequisites: INFO 701 [Min Grade: C]

INFO 704. Next-generation Sequencing Data Analysis. 3 Hours.
This course is aimed to equip participants with the essential knowledge and skills required to begin analyzing next-generation sequencing data and carry out some of the most common types of analysis. The topics covered in-depth during this course are the analysis of RNA-Seq, ChIP-Seq data, ATACseq data, and Single-cell data, with an optional Variant Calling session. The sessions will also include Introduction to next-generation sequencing (NGS) technologies, common NGS data analysis issues, applications of sequencing technologies, introduction to bioinformatics file formats (e.g. FASTQ, bam, bed) and bioinformatics toolkits. At the end of this course, participants will have the expertise to perform these data analysis independently.
Prerequisites: INFO 701 [Min Grade: C]

INFO 710. Programming with Biological Data. 3 Hours.
This course is a 4-week session, class meets 4 days a week from 9:00am-11:30am. Class will meet at Tinsley Harrison Tower, room 144. This course provides students necessary bioinformatics programming and data skills using Linux, MySQL and Python. Emphasis will be placed on best practices for conduct of reproducible research. The focus will be on practical computing, programming and data management concepts that can be applied to bioinformatics and data analysis problems.

INFO 711. Intermediate Statistical Analysis I. 3 Hours.
Students will gain a thorough understanding of basic analysis methods, elementary concepts, statistical models and applications of probability, commonly used sampling distributions, parametric and non-parametric one and two sample tests, confidence intervals, applications of analysis of two-way contingency table data, simple linear regression, and simple analysis of variance. Students are taught to conduct the relevant analysis using current software such as the Statistical Analysis System (SAS).

INFO 712. Visual Analytics for Bioinformatics. 3 Hours.
In this course, we will explore the use of visualization techniques as a concise and effective way to help analyze, understand, interpret and communicate complex biological data. Principles of design, visual rhetoric/communication, and appropriate usage will be introduced. We will cover representation of different data types, concentrating on those generated by data-rich platforms such as next-generation sequencing applications, flow/mass cytometry, and proteomics, and will discuss the use of visualization techniques applied to assessing data quality and troubleshooting. Various topics including dimension reduction, hierarchical visualizations, unsupervised learning, graph theory, networks/layouts and interactivity will be discussed. We will review the algorithmic underpinnings of various methods that lead to their appropriate and effective use. Finally, we will review a variety of genomics/bioinformatics-related visualization tools that are available. We will use Matlab throughout the course to create beautiful and effective visualizations.

INFO 751. Systems Biomedicine of Human Microbiota. 3 Hours.
The human microbiota is the collection of microorganisms (bacteria, archaea, fungi and viruses) that reside within human tissues and biofluids. Such resident microorganisms compose the majority of cells in human bodies and are key contributors to human development, health, and disease. However, most studies focus on genomics and microbiome statistical representations alone, while spatial-temporal analysis, multi-source data integration and modeling are necessary to predict and understand interactions between microorganisms, human hosts, and the environment. This course will highlight state-of-the-art microbiome/microbiota research and provide essential training in mathematical, computational and systems biology to derive integrative and predictive models of microbiota-host interactions in the context of human health and disease.
Prerequisites: INFO 701 [Min Grade: C] and (MA 560 [Min Grade: C] or BME 670 [Min Grade: C])

INFO 762. Biomedical Applications of Natural Language Processing. 3 Hours.
Students will be introduced to Natural Language Processing (NLP) including core linguistic tasks such as tokenization, lemmatization/stemming, POS tagging, parsing and chunking. Applications will focus on Deep Learning methods using pytorch with a focus in information extraction including Named Entity Recognition, semantic role labeling, word sense disambiguation, normalization, summarization, question answering and text classification. Applications and data will have a biomedical focus, but no biology or medical background is required.

INFO 773. Clinical Informatics Journal Club. 1 Hour.
Students will learn how to read, present, and critique research publications in clinical informatics. Journal Club participants will present high-impact recent journal publications selected by course instructors and learn how to read the paper, write critiques, and organize analysis insights into review papers. Students must show evidence of prior preparation prior to journal clubs and write critiques to show comprehension of the topics throughout the semester.

INFO 780. Implementation and Evaluation of Clinical Systems. 3 Hours.
Health information technology (HIT) tools such as Electronic Health Records (EHRs) are used to facilitate management of patient care data, to computerize clinical workflows, and to support health professionals in their medical decision making process. As a result of the U.S. Federal Government incentive program known as Meaningful Use, EHRs have been adopted on a national scale and are now used in almost every health care organization across the country. Although the literature exploring the impact of HIT adoption and use has also increased, previous studies have produced mixed results, leaving unanswered questions as to the impact of HIT on quality of care, patient safety, and health care providers' productivity. In this course, students will be introduced to project management tools and techniques commonly used for managing implementation of HIT systems as well as research approaches to conduct systematic evaluations of the impact of these systems on health care outcomes and organizations. This foundational course is intended for informatics majors and students in allied fields (e.g., health, biological, or computer sciences) who are interested in exploring implementation methods applicable to HIT systems such as EHRs and their components, as well as quantitative, qualitative, and mixed-methods approaches to conduct evaluations of HIT adoption and use. It is primarily intended for students who will pursue research careers in biomedical informatics and is the third course in a three-part series.
Prerequisites: INFO 797 [Min Grade: C]

INFO 790. Data Mining & Statistical Learning. 3 Hours.
Students will learn to discover and implement meaningful insights and knowledge from data. This course covers major concepts and algorithms of data mining. The course will be taught using the SAS Enterprise Miner program. The final project will demonstrate all the data mining techniques covered in the course and furthermore expose students working with real data. At the end of the course students will be proficient in utilizing data mining techniques to exploit data patterns and behavior, gain insider understanding of the data, and produce new knowledge that healthcare decision-makers can act upon. Furthermore, SAS Certified Predictive Modeler certification exam will be offered at the end of the course. Instructor permission is required.

INFO 791. Bioinformatics Seminar I. 1 Hour.
For doctoral student only. Students will learn how to prepare, present, and critique research presentations in bioinformatics by attending seminar presentations made by presenters. Seminars are presented by graduate students, faculty, visitors, or online speakers. Students must show evidence of prior preparation, active participation, and documented comprehension of the topics.
Prerequisites: INFO 701 [Min Grade: C]

INFO 792. Bioinformatics Seminar II. 1 Hour.
For doctoral student only. Students will learn how to prepare, present, and critique research presentations in bioinformatics by attending seminar presentations made by presenters. Seminars are presented by graduate students, faculty, visitors, or online speakers. Students must show evidence of prior preparation, active participation, and documented comprehension of the topics.
Prerequisites: INFO 791 [Min Grade: P]

INFO 793. Bioinformatics Journal Club. 2 Hours.
Students will learn how to read, present, and critique primary research publications in bioinformatics. Journal club participants will present high-impact recent journal publications selected by course instructors and learn how to read the paper, write critiques, and organize analysis insights into review papers. Students must show evidence of prior preparation prior to journal clubs and write critiques to show comprehension of the topics throughout the semester.

INFO 794. Advanced Bioinformatics Journal Club. 2 Hours.
Students will learn how to read, present, and critique primary research publications in bioinformatics. Journal club participants will present high-impact recent journal publications selected by course instructors and learn how to read the paper, write critiques, and organize analysis insights into review papers. Students must show evidence of prior preparation prior to journal clubs and write critiques to show comprehension of the topics throughout the semester.
Prerequisites: INFO 793 [Min Grade: P]

INFO 795. Special Topics in Bioinformatics. 3 Hours.
Topics of current research interest, such as metagenomics, microbiome, computational medicine, complex systems, deep learning in biology, artificial intelligence in biomedical, and translational bioinformatics applications. May be repeated as different sections taught by different instructors for credit. Permission of instructor is required.

INFO 796. Biomedical Informatics Methods I. 3 Hours.
Biomedical informatics is the art and science of collecting, representing and analyzing patient and biomedical information and translating insights from the information into better health and new medical discoveries. The spectrum of informatics applications ranges from molecules (bioinformatics) to individuals and populations (clinical and public health informatics). We will examine the scientific field and research methods that form the foundation for biomedical informatics research. The course will include didactics, readings, hands-on tool explorations, and a summative work product. This foundational course is intended for informatics majors and students in allied fields (e.g., health, biological, or computer sciences) who are interested in exploring the field of informatics.

INFO 797. Biomedical Informatics Methods II. 3 Hours.
Biomedical informatics is the art and science of collecting, representing and analyzing patient and biomedical information and translating insights from the information into better health and new medical discoveries. The spectrum of informatics applications ranges from molecules (bioinformatics) to individuals and populations (clinical and public health informatics). We will examine the scientific field and research methods that form the foundation for biomedical informatics research. The course will include didactics, readings, and applications in applying research methods, culminating in a research plan in grant proposal format and review by a mock panel. This foundational course is intended for informatics majors and students in allied fields (e.g., health, biological, or computer sciences) who are interested in exploring the field of informatics. It is primarily intended for students who will pursue research careers in biomedical informatics and is the second course in a two-part series.
Prerequisites: INFO 795 [Min Grade: B] or INFO 796 [Min Grade: B]

INFO 799. Bioinformatics Research for Dissertation. 1-12 Hour.
Admission to candidacy is required.