

Informatics, Data Science, and Biostatistics

Courses

BDSAI 5001 R for Biomedical Sciences

The course delves into the essential tools of Python programming within the context of biomedical informatics, biostatistics, and data science. This course emphasizes foundational programming skills in Python, crucial for data analysis in biomedical sciences. Participants will learn to set up and utilize the Jupyter environment for Python programming. The curriculum covers fundamental programming concepts, debugging techniques, and data manipulation using Pandas in Python. Practical sessions include working with diverse data formats such as CSV, JSON, and Excel, and handling data operations like data selection, filtering, grouping, and summarizing. By the end of the course, participants will be proficient in manipulating, analyzing, and managing, preparing you for more advanced studies or professional applications in biomedical informatics, biostatistics, and data analysis. Credit 1 unit.

Typical periods offered: Fall

BDSAI 5002 Python for Biomedical Sciences

The course delves into the essential tools of R programming within the context of biomedical informatics, biostatistics, and data science. This course emphasizes foundational programming skills in R, crucial for data analysis in biomedical sciences. Participants will learn to set up and utilize R Studio for R programming. The curriculum covers fundamental programming concepts, debugging techniques, and data manipulation using dplyr in R. Practical sessions include working with diverse data formats such as CSV, JSON, and Excel, and handling data operations like data selection, filtering, grouping, and summarizing. By the end of the course, participants will be proficient in manipulating, analyzing, and managing, preparing you for more advanced studies or professional applications in biomedical informatics, biostatistics, and data analysis. Credit 1 unit.

Typical periods offered: Fall Intersession

BDSAI 5003 Introduction to Biomedical Data Science

Introduction to Biomedical Data Science will provide students with an introduction to tools, theories and methods related to data modeling, management and query, data cleaning and analysis, and visualization that serve as the foundations for advanced topics in Biomedical Informatics and Data Science. The course consists of didactic lectures and experiential learning opportunities including hands-on laboratory sessions and a culminating project. Prior participation in "R and Python for Biomedical Sciences" course or completing a "Python Proficiency Test" are required, no other assumptions are made about computer science or clinical background. Credit 2 units.

Typical periods offered: Fall

BDSAI 5004 Introduction to Biomedical Informatics

This course offers an introduction to the definitions, theories, and methods that are foundational to Biomedical Informatics. Course content covers bioinformatics, clinical informatics, clinical research informatics, population health informatics, and imaging informatics.

Students will be introduced to topics such as data standards, clinical decision support, natural language processing, and data visualization. By the end of the course, participants will have a broad understanding of biomedical informatics and its applications, preparing them for further study or professional opportunities in the field.

Credit 2 units.

Typical periods offered: Fall

BDSAI 5005 Fundamentals of Biostatistics

This course is designed for students who want to develop a working knowledge of basic methods in biostatistics. The course is focused on biostatistical and epidemiological concepts and on practical hints and hands-on approaches to data analysis rather than on details of the theoretical methods. We will cover basic concepts in hypothesis testing, will introduce students to several of the most widely used probability distributions, and will discuss classical statistical methods that include t-tests, chi-square tests, regression analysis, and analysis of variance. Both in-class examples and homework assignments will involve extensive use of R.

Credit 2 units.

Typical periods offered: Fall

BDSAI 5111 Data Visualization

This course introduces the fundamental principles and practical applications of data visualization in biomedical data science. Students will learn how to translate complex biomedical and clinical data into clear and interpretable visual representations, and apply visualization techniques to explore, analyze, and interpret biomedical and clinical datasets. The course covers foundational concepts, such as perception, color theory, and visual design, along with best practices in data communication. Students will gain hands-on experience with visualization tools to create static and interactive visualizations. The student will apply visualization techniques to real-world biomedical datasets, integrating visual analytics into workflows, and communicating data insights to diverse research and clinical audiences.

Credit 3 units.

Typical periods offered: Fall

BDSAI 5121 Electronic Health Records: Foundations

A course covering healthcare information technology (health IT), specifically focusing on electronic health records (EHRs). This course is designed to help students understand structures, functions, roles, and other factors that influence the use and impact of EHRs. It also addresses current issues and strategies related to health IT.

Credit 3 units.

Typical periods offered: Fall, Spring

BDSAI 5121 Electronic Health Records: Foundations

A course covering healthcare information technology (health IT), specifically focusing on electronic health records (EHRs). This course is designed to help students understand structures, functions, roles, and other factors that influence the use and impact of EHRs. It also addresses current issues and strategies related to health IT.

Credit 3 units.

Typical periods offered: Fall, Spring

BDSAI 5123 Advanced Multi-Omic Modeling and Systems Analysis

This advanced course addresses the computational and statistical challenges of integrating diverse high-throughput biological data—including genomics, transcriptomics, proteomics, metabolomics, and epigenomics. The curriculum covers data harmonization, preprocessing techniques, and advanced statistical methods such as matrix factorization, kernel-based techniques, and Bayesian inference,

alongside application of network and systems biology approaches for predictive modeling, biomarker discovery, and clinical translation in personalized medicine. Through computational labs, literature reviews, and a semester-long project, MS and PhD students with strong backgrounds in statistics and programming (R or Python) will develop the ability to critically evaluate, design, and implement robust multi-omic data analysis pipelines, equipping them with essential skills for cutting-edge biological research and clinical discovery. This advanced, doctoral-level course provides a deep dive into the computational and statistical challenges inherent in integrating diverse high-throughput biological data (genomics, transcriptomics, proteomics, metabolomics, and epigenomics). The course begins by reviewing state-of-the-art data harmonization and preprocessing techniques necessary for combining datasets of varying dimensionality and noise profiles. Key themes include advanced statistical modeling for integration, such as matrix factorization methods, kernel-based techniques, and Bayesian inference, alongside the application of network and systems biology approaches to contextualize findings. Students will explore predictive modeling, biomarker discovery, and clinical translation within the framework of personalized medicine. The target audience is MS and PhD students in Biomedical Informatics, Computational Biology, and related quantitative disciplines who possess a strong foundation in statistics and programming (R or Python). Through computational labs, critical literature reviews, and a semester-long, project-based assessment, students will transition from theoretical knowledge to practical application. Upon completion, students will be able to critically evaluate, design, and implement robust multi-omic data analysis pipelines, equipping them with essential skills for cutting-edge biological research and clinical discovery.

Credit 3 units.

Typical periods offered: Fall

BDSAI 5131 Survival Analysis

This course introduces the fundamental theoretical and applied principles of survival analysis for modeling time-to-event data. Core topics include survival and hazard functions, censoring and truncation, Kaplan–Meier and Nelson–Aalen estimators, cohort life tables, and likelihood construction for censored and truncated data. The course covers estimation of hazard and survival functions, the Cox proportional hazards (PH) model with fixed and time-dependent covariates, and methods for model selection. Additional topics include regression diagnostics for survival models, stratified PH models, parametric regression models, and competing risks analysis. Computer lab sessions provide hands-on experience with real-world datasets to reinforce methodological concepts and data analysis skills.

Credit 3 units.

Typical periods offered: Fall

BDSAI 5213 Causal Artificial Intelligence for Healthcare

This course provides an introduction to causal artificial intelligence for healthcare, focusing on how causal reasoning and modern machine learning methods can be used to support reliable decision-making from real-world health data. Students will learn how to move beyond prediction toward answering clinically meaningful “what if” questions, such as estimating treatment effects, understanding sources of bias, and evaluating model robustness across populations and settings. Topics include randomized and observational study designs, causal diagrams, target trial emulation, counterfactual reasoning, heterogeneous treatment effects, invariance and transportability, algorithmic fairness, and emerging approaches to causal representation learning. Throughout the course, methodological concepts are motivated by real clinical and public-health use cases.

Credit 3 units.

Typical periods offered: Fall, Spring

BDSAI 5214 Entrepreneurship in Healthcare Informatics Using Data Science

This course introduces students to entrepreneurship within the context of healthcare informatics and data science. Students will explore how data-driven technologies such as AI, machine learning, electronic health records, and digital health tools can be translated into viable healthcare solutions. The course emphasizes identifying real-world healthcare problems, evaluating market opportunities, and navigating the legal, ethical, and regulatory landscape of healthcare innovation. Through case studies, hands-on activities, and a team-based project, students will develop and pitch a healthcare informatics startup concept grounded in clinical, technical, and business considerations.

Credit 3 units.

Typical periods offered: Fall, Spring

BDSAI 5215 Implementation of Health Information Technology into Real-World Settings

This course will serve as an introduction to implementation science and its use in implementing and evaluating health information technology into real-world settings. Students will learn the foundations of implementation science including how to design, conduct, and evaluate an implementation project. Key conceptual topics to be covered include implementation readiness, strategies, and outcomes, and key methodological topics to be covered include qualitative focus groups and interviews, engaging and managing advisory or stakeholder boards, and administering surveys to assess implementation outcomes.

Credit 3 units.

Typical periods offered: Fall, Spring

BDSAI 5216 Modern Data and AI Architectures for Biomedical Applications

Modern Data and AI Architectures for Biomedical Applications explores the design, implementation, and evaluation of scalable data and artificial intelligence architectures used in contemporary biomedical research and healthcare systems. The course covers end-to-end pipelines—from data acquisition and integration to model deployment—focusing on structured and unstructured biomedical data such as clinical records, omics data, and medical imaging. Emphasis is placed on modern architectural paradigms, including cloud-native systems, distributed computing, data lakes, and MLOps, alongside considerations of data governance, privacy, security, and regulatory compliance in biomedical contexts.

Credit 3 units.

Typical periods offered: Fall, Spring

BDSAI 5301 Seminar Series for Biomedical Sciences

This seminar series introduces students to research and emerging trends in biomedical informatics, data science, and biostatistics. Through presentations from invited speakers, students will learn about ongoing research initiatives and gain insight into career paths and types of interdisciplinary collaborations. Students will gain exposure to different scientific communication styles and develop critical evaluation skills through post-seminar discussions.

Credit 1 unit.

Typical periods offered: Fall

**BDSAI 5302 Journal Club for Biomedical Sciences**

Guided journal club discussions will build skills in critically appraising research articles, enabling students to evaluate methodological rigor, interpret findings, and consider their impact on science and practice. Students will gain experience in both leading discussions and providing constructive peer feedback on scientific presentations. The course is designed to broaden students' perspectives, strengthen their analytical skills, and support their professional development.

Credit 1 unit.

Typical periods offered: Spring

BDSAI 5901 Biomedical Data Science and AI Practicum

Students in the Internship course will acquire critical professional experience so that they will be well prepared to enter the scientific workforce upon graduation. This course provides a guided, mentored opportunity for students to build marketable skills, experience hands on applied work on open ended problems outside of the classroom, bolster professional and scientific communication skills, and gain greater understanding of their professional interests in aspects of biostatistics or biomedical informatics. Students will have an opportunity to work with experienced mentors (PIs) on a range of projects that may include data management, data analysis, study design, protocol development, and potentially, publication of scientific manuscripts. As part of the Internship requirements will give a presentation of the Internship experience. The grade (pass/fail) for each student will be determined in consultation with the mentor.

Credit 3 units.

Typical periods offered: Fall, Spring

BDSAI 5902 Biomedical Data Science and AI Mentored Research

Students will demonstrate how to synthesize and apply the full spectrum of biomedical informatics or biostatistics theories and methods included in the program curriculum. The mentored research project requires students to formulate research questions that focus on the development or extension of a theoretical framework or a novel method with relevance to the field of informatics or biostatistics, resulting in a report that outlines the student's topic selection and the design, conduct, and results of the student's research. Each trainee will also be expected to present their project and its outcomes or findings in a public seminar, where questions will be posed by both the audience and a committee of faculty members. The specific selection of the internship or mentored research project track as part of a trainee's degree program is to be discussed with and approved by the individual's faculty and academic adviser. The course instructor will determine the grade (pass/fail) in consultation with the mentors.

Credit 3 units.

Typical periods offered: Fall, Spring

BDSAI 7883 Master's Continuing Student Status

Full-time graduate research

Credit 0 units.

Typical periods offered: Fall, Spring, Summer
