Ashlynn Fuccello
Hendrix College
Class of 2021

Interdisciplinary Studies Major in Bioinformatics

Bioinformatics is an interdisciplinary field of study that applies the analytical tools from mathematics and computer science to the analysis of biological data. This field of study is rapidly growing as mass amounts of biological data are being accessed through genome sequencing and other advancements in research and the previous systems for analyzing biological data can't keep up. The skills learned in computer science and math classes allow for large-scale, data intensive biological problems to be attacked from a computational perspective.

Bioinformatics is a heavily research oriented field that requires a deep understanding of both biological data and computational processes. A graduate level degree is necessary in order to pursue a career within the field and so this major has been designed to prepare any student to gain admission into graduate programs in bioinformatics. Course requirements for graduate programs at John Hopkins, University of Arkansas at Little Rock, and Carnegie Melon University, as well as undergraduate programs in bioinformatics from University of California at Los Angeles and the University of Maryland, were considered when building the following major. The main take away from these majors was to place focus on genetics and molecular biology courses, algorithms and database courses, and applied mathematics classes. The one constant across every single one of the programs was the requirement of a strong applied statistics class. Since the probability and statistics class at Hendrix has less focus on applied statistics, this major includes the option to take an applied statistics class at a university in Spain. The course syllabus for this statistics class is attached.

A project following the guidelines of those in the computer science senior capstone has been chosen for this major. This is a final research project that is student designed and led

meaning that the project can be bioinformatics based and still follow the guidelines of the senior capstone because any bioinformatics research project involves the use of computer science skills. Since the senior seminar class is not a part of the capstone for this major, the project will be extended to be a year-long project as opposed to only a semester.

All of the following:

CSCI 150	Foundations of Computer Science
CSCI 151	Data Structures
CSCI 382	Algorithms and Prob Solving Paradigms
BIOL 150	Cell Biology
BIOL 150	Genetics
MATH 130	Calculus I
MATH 140	Calculus II
MATH 240	Discrete Mathematics

Either MATH 310 Probability and Statistics or Statistics at Universidad Carlos IIII de Madrid

At least 2 of the following:

BIOL 355	Advanced Cell Biology
BIOL 470	Advanced Genetics
BIOL 465	Molecular Evolution and Bioinformatics
BIOL 310	Developmental Biology
BIOL 460	Evolution
BIOL 325	Cellular and Molecular Neuroscience
BIOL 430	Immunology
CSCI 340	Database and Web Systems

At least 1 of the following:

MATH 260	Introduction to Advanced Mathematics
MATH 365	Mathematical Models
MATH 270	Linear Algebra
MATH 260	Differential Equations
MATH 340	Combinatorics

SENIOR CAPSTONE: An individual project following the guidelines and structure of the CSCI capstone.

Syllabus for Statistics at Universidad Carlos III de Madrid

BLOCK I: PROBABILITY

1. Introduction to Probability
1.1 Introduction
1.2 Random phenomena
1.3 Definition of probability and properties
1.4 Conditional probability
1.5 Bayes Theorem

2. Random variables
2.1 Definition of random variable
2.2 Discrete random variables
2.3 Continuous random variables
2.4 Characteristic features of a random variable
2.5 Transformations of random variables
2.5 Random vectors

3. Distribution models
3.1 Binomial distribution
3.2 Poisson distribution
3.3 Uniform distribution (continuous)
3.4 Exponential distribution
3.5 Normal distribution (with CLT)
3.6 Multivariate normal distribution

BLOCK II: STATISTICS

4. Parameter Estimation
4.1 Introduction and basic concepts
4.2 Sampling distributions
4.3 Maximum Likelihood Estimation
4.4 Properties of Maximum Likelihood Estimators
4.5 Inference based on MLEs

5. Statistical Inference
5.1 Introduction
5.2 Estimators and their distributions
5.3 Confidence Intervals
5.4 Hypothesis testing
5.5 Particular tests on a single sample
5.6 Comparison of two populations

BLOCK III: REGRESSION

6. Linear regression
6.1 Introduction
6.2 Simple linear regression
6.3 Multiple linear regression

Approval Signatures

Dr. Brent Yorgey

Dr. George Harper

Dr. David Sutherland